

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 14:33:42 ; Search time 3672.39 Seconds

(without alignments)
11294.891 Million cell updates/sec

Title: US-10-076-747-7

Perfect score: 957

Sequence: 1 ttggagcttaccatcatcatg.....gacagagagaccagagct 957

Scoring table: IDENTITY NJC

Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```
Database :
1: gb_ba:*
2: gb_ba:*
3: gb_ba:*
4: gb_ba:*
5: gb_ba:*
6: gb_ba:*
7: gb_ba:*
8: gb_ba:*
9: gb_ba:*
10: gb_ba:*
11: gb_ba:*
12: gb_ba:*
13: gb_ba:*
14: gb_ba:*
15: gb_ba:*
16: gb_ba:*
17: gb_ba:*
18: gb_ba:*
19: gb_ba:*
20: gb_ba:*
21: gb_ba:*
22: gb_ba:*
23: gb_ba:*
24: gb_ba:*
25: gb_ba:*
26: gb_ba:*
27: gb_ba:*
28: gb_ba:*
29: gb_ba:*
30: gb_ba:*
31: gb_ba:*
32: gb_ba:*
33: gb_ba:*
34: gb_ba:*
35: gb_ba:*
36: gb_ba:*
37: gb_ba:*
38: gb_ba:*
39: gb_ba:*
40: gb_ba:*
41: gb_ba:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56.4	5.9	991	11	PM12H12B
2	53.8	5.6	7218	6	1664455
3	53.4	5.5	218201	2	AC141731
4	53	5.5	1392	11	PM12B10G
5	53	5.5	160634	2	AC141999
6	53	5.5	177485	2	AC090975
7	50.8	5.3	10732	6	E32986
8	50.8	5.3	125020	2	AF429315
9	50.4	5.3	123570	2	AC142034
10	50.4	5.3	125020	9	AF429315
11	50.2	5.2	46251	9	AD000671
12	50.2	5.2	194835	2	AC022315
13	50	5.2	152666	2	AC143344
14	49.6	5.2	2000	6	AX655393
15	49.6	5.2	142852	8	AF151259
16	49.6	5.2	163032	2	AC144742
17	49.6	5.2	187223	2	AC008333
18	49.4	5.2	1393	11	PM1H12G
19	49.2	5.1	189230	2	AC120836
20	49	5.1	177877	2	AC133243
21	48.8	5.1	85434	2	AC066610
22	48.4	5.1	1070	9	ORARGIT
23	48.2	5.0	166352	9	AC068530
24	48.2	5.0	177055	9	AC090360
25	48.2	5.0	276546	9	AC139100
26	47.6	5.0	150942	2	AC130463
27	47.4	5.0	576	9	HS4335249
28	47	4.9	72847	2	AC023563
29	47	4.9	197082	9	AC006249
30	47	4.9	239130	2	AC079420
31	47	4.9	256673	2	AC087146
32	46.8	4.9	140176	2	AC016411
33	46.8	4.9	160587	8	OSUN00222
34	46.6	4.9	39954	9	AP006257
35	46.4	4.8	125402	9	AC114815
36	46.4	4.8	136551	2	AC048354
37	46.4	4.8	202058	2	AC016913
38	46	4.8	1065	11	PM2B12B
39	46	4.8	63082	2	AC022663
40	46	4.8	64189	2	AC015511
41	46	4.8	110000	2	AC143301_0
42	46	4.8	156289	2	AC144155
43	46	4.8	262074	2	AC098624
44	46	4.8	272166	2	AC120907
45	45.8	4.8	83569	2	AC022967

ALIGNMENTS

RESULT 1
PM12H12B
LOCUS Penicillium marneffei STS, clone pm12h12.b, sequence tagged site.
DEFINITION
ACCESSION AF684455
VERSION AF684455.1 GI:19338056
KEYWORDS
SOURCE
ORGANISM
STS
Penicillium marneffei
Penicillium marneffei
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
REFERENCE
1 Yuen, K.Y., Pascual, G., Wong, S.S., Glaeser, P., Woo, P.C., Kunst, F.,
Cal, J.C., Cheung, E.Y., Medigue, C. and Danchin, A.
Exploring the Penicillium marneffei genome

JOURNAL Arch. Microbiol. 179 (5), 339-353 (2003)
 MEDLINE 22595073
 PUBMED 12640520
 REFERENCE 2 (bases 1 to 991)
 AUTHORS Danchin, A. and Pascal, G.
 TITLE Direct Submission
 JOURNAL Submitted (08-MAR-2002) Danchin A., HKU-Pasteur Research Centre,
 Dexter HC Man Building 8, Sarssoon Road, Pokfulam, Hong Kong

FEATURES
 Source
 1. .991
 /organism="Penicillium marneffei"
 /mol_type="genomic DNA"
 /db_xref="taxon:37727"
 /clone="pm12h12.b"

ORIGIN
 Query Match 5.3%; Score 56.4; DB 11; Length 991;
 Best Local Similarity 50.2%; Pred. No. 0.039;
 Matches 135; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 685 CCGGAGACGGGCGGAGACCCCGGGGAGAGAGCCCACTGGGGAGGGGAGGAG 744
 Db 480 CGGGGCCCCCGCGCGCGGGGGGGCGCCCGGGGCGGGGGGGCGCGGGGGG 539
 QY 745 ACAAGATACGAGAGCGGAGCCACGCGCCCGTCATGAGACGAGAGCGCCGAGGGA 804
 Db 540 GCGGGGGCGCGCGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGGGG 599
 QY 805 GCGGACAGCCGAGAGCGCGCGCCCAAGAGCGAGCGCCCAAGCGCCCTAGGCGCTG 864
 Db 600 CGCGCGCGCGCGCGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGG 659
 QY 865 CGCCAGCGCGCGCGCGCGCGCGCGAGAGCGGGGGCGAGACCGGAGAGACCGACCGAGC 924
 Db 660 CGGGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGG 719
 QY 925 CCGACCGGGCGGCGGAGCGAGCGACCGC 953
 Db 720 CCGCGCGGG 748

RESULT 2
 166494/c
 LOCUS 166494 7218 bp DNA linear PAT 28-DEC-1997
 DEFINITION Sequence 14 from patent US 5670367.
 ACCESSION 166494
 VERSION 166494.1 GI:2724471
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 7218)
 AUTHORS Dornier, F., Scheiflinger, F. and Falkner, F. Gunter.
 TITLE Recombinant fowlpox virus
 JOURNAL Patent: US 5670367-A 14 23-SEP-1997;
 FEATURES
 Location/Qualifiers
 1. .7218
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN

Query Match 5.6%; Score 53.8; DB 6; Length 7218;
 Best Local Similarity 1.9%; Pred. No. 0.13; Mismatches 141; Indels 0; Gaps 0;
 Matches 7; Conservative 219; Mismatches 141; Indels 0; Gaps 0;

QY 581 ATGGCGTACCGGCGCATCAATTGCGCGGAGCTACAGACTTAAAGTTACCGCGCA 640
 Db 1443 ATTGGTACGRR 1384
 QY 641 GCGGCGAAGACAGCGATTTCAGCGAGTCCCAACAGAGAGGGCGGCGAGCGAGCG 700
 Db 1383 RRR 1324

QY 701 GACCCGCGGCGGAGACCCCACTGCGGCGAGCGGCGAGCGAGACAGATACGAGCG 760
 Db 1323 RRR 1264
 QY 761 GAGCCACCGCGGCGCCGTCATGAGACGAGCGCGCGGAGAGCGAGACCGGAGCA 820
 Db 1263 RRR 1204
 QY 821 GCGGCGAAGACGAGCGCGGCGAAGCGAGCGCGCTAGGCGCTGCGGCGAGCGCG 880
 Db 1203 RRR 1144
 QY 881 GCGGAGAGCGGCGGCGGAGACCGCGAGACCGGAGACCGGAGCGCGCGCGCGG 940
 Db 1143 RRR 1084
 QY 941 AGCAGAG 947
 Db 1083 RRRRRRR 1077

RESULT 3
 AC141731
 LOCUS Apis mellifera clone CH224-58F2, WORKING DRAFT SEQUENCE, 91
 DEFINITION unordered pieces.
 ACCESSION AC141731
 VERSION AC141731.1 GI:29123915
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
 SOURCE Apis mellifera (honeybee)
 ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea; Apoidea; Apis.
 1 (bases 1 to 218201)
 Muzny, D.M., Adams, C., Adio-Oduola, B., All-ouman, F.R., Allen, C., Alsbrooks, S.L., Amaralunga, H.C., Are, O.R., Ayale, M., Benke, T., Barbarella, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, R., Dugan-Rocha, S., Durbin, K.J., Eamhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Hulik, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Koryah, J., Kovar, C., Kratovic, J., Kureishi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtharge, O., Lien, C., Liu, J., Liu, W., Louleghed, H., Lorado, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawiney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokemko, S., Ogih, M., Okumori, G., Oragunye, N., Oyedele, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Sickens, R., Primm, E., Pu, L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojurokan, I., Rolfe, M., Ruiz, S., Severy, G., Scherier, S., Scott, G., Shen, H., Shoshitaishvili, N., Stasom, I., Sodergren, E., Sonakke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, S., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umanai, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Mu C., Mu Y., Mu Y.F., Zhou J., Zorrilla S., Nelson D.,
Weinstock G., and Gibbs R.
Direct Submission
Unpublished
2 (bases 1 to 218201)
Worley, K.C.
Submitted (19-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: AMEV
Center clone name: CH224-58F22
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990129
Consensus quality: 195857 bases at least Q40
Consensus quality: 207269 bases at least Q30
Consensus quality: 213633 bases at least Q20
Estimated insert size: 183061; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 91 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1459: contig of 1459 bp in length
* 1460 1559: gap of unknown length
* 1560 2927: contig of 1368 bp in length
* 2928 3027: gap of unknown length
* 3028 4064: contig of 1037 bp in length
* 4065 4164: gap of unknown length
* 4165 5450: contig of 1286 bp in length
* 5451 5550: gap of unknown length
* 5551 6910: contig of 1360 bp in length
* 6911 7010: gap of unknown length
* 7011 8455: contig of 1445 bp in length
* 8456 8555: gap of unknown length
* 8556 9628: contig of 1073 bp in length
* 9629 9728: gap of unknown length
* 9729 10754: contig of 1026 bp in length
* 10755 10854: gap of unknown length
* 10855 12039: contig of 1185 bp in length
* 12040 12139: gap of unknown length
* 12140 13142: contig of 1003 bp in length
* 13143 13242: gap of unknown length
* 13243 14333: contig of 1091 bp in length
* 14334 14433: gap of unknown length
* 14433 15532: contig of 1099 bp in length
* 15533 15632: gap of unknown length
* 15633 16838: contig of 1206 bp in length
* 16839 16938: gap of unknown length
* 16939 18146: contig of 1208 bp in length
* 18147 18246: gap of unknown length
* 18247 19466: contig of 1220 bp in length
* 19467 19566: gap of unknown length
* 19567 20905: contig of 1339 bp in length
* 20906 21005: gap of unknown length
* 21006 22525: contig of 1520 bp in length
* 22526 22625: gap of unknown length
* 22626 23739: contig of 1114 bp in length

23740 23840 23839: gap of unknown length
* 23840 24845: contig of 1006 bp in length
* 24846 24945: gap of unknown length
* 24946 26300: contig of 1355 bp in length
* 26301 26400: gap of unknown length
* 26401 27643: contig of 1243 bp in length
* 27644 27743: gap of unknown length
* 27744 28940: contig of 1197 bp in length
* 28941 29040: gap of unknown length
* 29041 30245: contig of 1205 bp in length
* 30246 30345: gap of unknown length
* 30346 31518: contig of 1173 bp in length
* 31519 31618: gap of unknown length
* 31619 33173: contig of 1555 bp in length
* 33174 33273: gap of unknown length
* 33274 34738: contig of 1465 bp in length
* 34739 34838: gap of unknown length
* 34839 35982: contig of 1144 bp in length
* 35983 36082: gap of unknown length
* 36083 37444: contig of 1362 bp in length
* 37445 37544: gap of unknown length
* 37545 38752: contig of 1208 bp in length
* 38753 38852: gap of unknown length
* 38853 40154: contig of 1302 bp in length
* 40155 40254: gap of unknown length
* 40255 42036: contig of 1782 bp in length
* 42037 42137: gap of unknown length
* 42138 43361: contig of 1225 bp in length
* 43362 43461: gap of unknown length
* 43462 44861: contig of 1400 bp in length
* 44862 44961: gap of unknown length
* 44962 46359: contig of 1398 bp in length
* 46360 46459: gap of unknown length
* 46460 47564: contig of 1105 bp in length
* 47565 47664: gap of unknown length
* 47665 49513: contig of 1849 bp in length
* 49514 49613: gap of unknown length
* 49614 50721: contig of 1108 bp in length
* 50722 50821: gap of unknown length
* 50822 52553: contig of 1732 bp in length
* 52554 52653: gap of unknown length
* 52654 54428: contig of 1775 bp in length
* 54429 54528: gap of unknown length
* 54529 55799: contig of 1271 bp in length
* 55799 55899: gap of unknown length
* 55900 56993: contig of 1094 bp in length
* 56994 57093: gap of unknown length
* 57094 58796: contig of 1603 bp in length
* 58797 58896: gap of unknown length
* 58897 60055: contig of 1259 bp in length
* 60056 60155: gap of unknown length
* 60156 61586: contig of 1431 bp in length
* 61587 61686: gap of unknown length
* 61687 63051: contig of 1365 bp in length
* 63052 63151: gap of unknown length
* 63152 64599: contig of 1448 bp in length
* 64600 64699: gap of unknown length
* 64700 66585: contig of 1846 bp in length
* 66586 66645: gap of unknown length
* 66646 68608: contig of 1963 bp in length
* 68609 68708: gap of unknown length
* 68709 70013: contig of 1305 bp in length
* 70014 70113: gap of unknown length
* 70114 71424: contig of 1311 bp in length
* 71425 71524: gap of unknown length
* 71525 73106: contig of 1582 bp in length
* 73107 73206: gap of unknown length
* 73207 74473: contig of 1267 bp in length
* 74474 74573: gap of unknown length
* 74574 76548: contig of 2075 bp in length
* 76549 76749: gap of unknown length
* 76750 78054: contig of 1306 bp in length
* 78055 78154: gap of unknown length

*	78155	79466:	contig of 1312 bp in length
*	79467	79566:	gap of unknown length
*	79567	81155:	contig of 1589 bp in length
*	81156	81255:	gap of unknown length

Query Match	5.6%	Score 53.4;	DB 2;	Length 218201;
Best Local Similarity	47.6%;	Pred. No. 0.13;		
Matches 138; Conservative	0;	Mismatches 152;	Indels 0;	Gaps 0

Accession	Sequence	Length
QY	CGCAGCGGCCAAGAACAGCGTTTGCAGCGAATGCCAAACACAGAGGGCGCAGACCGGAC	696
637		
Db	CCCCGGGGGCACCAGGCGCAAGCCCAAGCAGCGAGGANNNGAGACAGCGCCGACCGAGNNGCCCA	16486
16483		

Accession	Gene	Length (bp)
Dy 697	AGGCGATCCCCGGGCGAGAGCCCCCACTGCGCAGGGCCAGGCGCAAGGACAGATTCAGAG	756
Db 164863	GAGCGCGCCCGCGAGGCCCCCGAGCGCGGAGCAGCAGAGCGAGAACGAGAAACAGGGGACCGAGC	164922

Db 164923 GCGAAGCGCAACCCCGCCCGAGGCGACGCGAGGAGCGCCCGCGACACNCGGGAAA 164982

Dy	817	AGCAGGCGCCGCAGACCGACCGGCAGGCCACACGGCGCTTAGGCGCTTGGCCCAAGCGGCGG	876
Dd	164983	GACCCCGCGGAAAGGGCACCGCCACGCMGACCCTCGCGCGGGAAACGGCGCTGAGAGNCG	165042

oy	877	CCACGCGGAGGGCGGGCGGAGCACCGCAGGAGACCAGCAACCGAACCC	926
Db	165043	CCGCGGCACAGCGGGCGGGCGGAGCACCGGGANCAACCCGCACGGGCC	165092

PM12B10G/c	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE
PM12B10G	1392 bp	DNA	linear	STS 29-MAY-2000				
		Penicillium marneffe	STS, clone pm12b10.g,					
			sequence tagged site.					
AL684308								
AL684308.1	GI:19338017							
		STS.						
		Penicillium marneffe						
		Penicillium marneffe						
		Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;						
		Eurotiates; Trichocomaceae; mitosporic Trichocomaceae; Penicillium						

REFERENCE
AUTHORS
TITLE
JOURNAL

Yen, K.Y., Pascal, G., Wong, S.S., Glaser, P., Woo, P. C., Kunst, F.,
Cai, J., Cheng, E.Y., Medigue, C. and Danchin, A.
Exploring the *Penicillium marneffei* genome
Arch. Microbiol. 179 (5), 339-353 (2003)

REFERENCE 2 (bases 1 to 1392)
AUTHORS Danchin, A. and Pascal, G.
TITLE Direct Submission
JOURNAL Submitted (08-MAR-2002) Danchin A., HKU-Pasteur Research Centre
DESCRIPTOR HC Man Building 8, Sassoon Road, Pokfulam, Hong Kong
LOCATION/Qualifiers

ORIGIN

Query Match	5.5%;	Score 53;	DB 11;	Length 1392;
Best Local Similarity	51.5%;	Pred. No. 0.22;		
Matches 139;	Conservative	0;	Mismatches 130;	Indels 1;
			Gaps	1

QY 683 GGCCTCAGACCGGCAAGCGGACCCCGGCGAGAGCCCTGCGGC - AGGCGAGGCGA 741

Dp 503 GCCCGAAGGCTCGCGCGCTGCGCGAGCGCGCGCGCGCGCGCGCGCG 444

DY 742 AGGACAGATTACGAGACGCCAGCCACACGCCCGCCCTGCATGAGACGGAGACCAGC 801

Dd 443 CCGGNCGCCGCGCGCCGNCGCCGACACACACAGGAGGCCACAGCGCGGGCCGCGCCGCG 384

QY	80	GAGGCCCAACCCGAAAGCAGCGCGCAAGACGCCCGCTAAGGC	861
Dd	383	GCGCGAATCCCGGAGCGCGCGCGCGAACCGCGCGCGCGCGCGCG	324
QY	862	CTAGCGCAAGCGCGCGCCACACGCGAGAAGCGCGCGAGCACCGACCA	921
Dd	323	GCGCGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	264
QY	922	GACCCGACCGCGCCGAGGCGACAGACGAC	951
Dd	263	GCGCGCGCGCGAGACTGACACACCGAGGC	234

RESULT 5	
AC141999/c	
LOCUS	AC141999
DEFINITION	AC141999 160634 bp DNA linear HTG 21-MAR-2003
	Rattus norvegicus clone CH230-270C13, *** SEQUENCING IN PROGRESS
ACCESSION	AC141999
VERSION	AC141999.1 GI:29135470
KEYWORDS	HTG; HTGS_PHASE1.
SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus

REFERENCE
AUTHORS
1 (bases 1 to 160634)
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Zhao S., Dunn D., von Niederhausern A., Weiss R., Smith D.R.,
 Holt R.A., Smith H.O., Weinstock G. and Gibbs R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 160634)
 Wozley K.C.
 Submitted (21-MAR-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GRBY
 Center clone name: CH230-270C13
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 126970 bases at least Q40
 Consensus quality: 134664 bases at least Q30
 Consensus quality: 140158 bases at least Q20
 Estimated insert size: 126396; sum-of-coverage estimation
 Quality coverage: 2x in Q20 bases; sum-of-coverage estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 61 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1002: contig of 1002 bp in length
 * 1003 1102: gap of unknown length
 * 1103 1104: contig of 1002 bp in length
 * 2105 2204: gap of unknown length
 * 2205 3741: contig of 1537 bp in length
 * 3742 3841: gap of unknown length
 * 3842 5455: contig of 1614 bp in length
 * 5456 5555: gap of unknown length
 * 5556 7319: contig of 1764 bp in length
 * 7320 7419: gap of unknown length
 * 7420 8494: contig of 1075 bp in length
 * 8495 8594: gap of unknown length
 * 8595 10312: contig of 1718 bp in length
 * 10313 10412: gap of unknown length
 * 10413 11608: contig of 1196 bp in length
 * 11609 11708: gap of unknown length
 * 11709 12796: contig of 1088 bp in length
 * 12797 12896: gap of unknown length
 * 12897 14011: contig of 1115 bp in length
 * 14012 14111: gap of unknown length
 * 14112 15372: contig of 1261 bp in length
 * 15373 15472: gap of unknown length
 * 15473 16702: contig of 1230 bp in length
 * 16703 16802: gap of unknown length
 * 16804 18540: contig of 1738 bp in length
 * 18541 18640: gap of unknown length
 * 18641 18848: contig of 1208 bp in length
 * 18849 19948: gap of unknown length
 * 19949 21085: contig of 1137 bp in length
 * 21086 21185: gap of unknown length
 * 21186 22538: contig of 1353 bp in length
 * 22539 22639: gap of unknown length
 * 22639 24199: contig of 1561 bp in length
 * 24200 24299: gap of unknown length
 * 24300 25490: contig of 1191 bp in length

* 25491 25590: gap of unknown length
 * 25591 26810: contig of 1220 bp in length
 * 26811 26910: gap of unknown length
 * 26911 28333: contig of 1423 bp in length
 * 28334 28433: gap of unknown length
 * 28434 29824: contig of 1391 bp in length
 * 29825 29924: gap of unknown length
 * 29925 31363: contig of 1439 bp in length
 * 31364 31463: gap of unknown length
 * 31464 33704: contig of 2241 bp in length
 * 33705 33804: gap of unknown length
 * 33805 35873: contig of 2069 bp in length
 * 35874 35973: gap of unknown length
 * 35974 38042: contig of 2069 bp in length
 * 38043 38142: gap of unknown length
 * 38143 39581: contig of 1439 bp in length
 * 39582 39681: gap of unknown length
 * 39682 41011: contig of 1330 bp in length
 * 41012 41111: gap of unknown length
 * 41112 42348: contig of 1237 bp in length
 * 42349 42448: gap of unknown length
 * 42449 44108: contig of 1660 bp in length
 * 44109 44208: gap of unknown length
 * 44209 47160: contig of 2952 bp in length
 * 47161 47260: gap of unknown length
 * 47261 49178: contig of 1918 bp in length
 * 49179 49278: gap of unknown length
 * 49279 51273: contig of 1995 bp in length
 * 51274 51373: gap of unknown length
 * 51374 53195: contig of 1822 bp in length
 * 53196 53295: gap of unknown length
 * 53296 55875: contig of 2580 bp in length
 * 55876 55975: gap of unknown length
 * 55976 58197: contig of 2222 bp in length
 * 58198 58297: gap of unknown length
 * 58298 60391: contig of 2094 bp in length
 * 60392 60491: gap of unknown length
 * 60492 61807: contig of 1316 bp in length
 * 61808 61907: gap of unknown length
 * 61909 64495: contig of 2588 bp in length
 * 64496 64595: gap of unknown length
 * 64596 67408: contig of 2813 bp in length
 * 67409 67508: gap of unknown length
 * 67509 69971: contig of 2463 bp in length
 * 69972 70071: gap of unknown length
 * 70072 72989: contig of 2918 bp in length
 * 72990 73089: gap of unknown length
 * 73090 74939: contig of 1850 bp in length
 * 74940 75039: gap of unknown length
 * 75040 76951: contig of 1912 bp in length
 * 76952 77051: gap of unknown length
 * 77052 80355: contig of 3204 bp in length
 * 80356 80355: gap of unknown length
 * 80356 82038: contig of 1683 bp in length
 * 82039 82138: gap of unknown length
 * 82139 85079: contig of 2941 bp in length
 * 85080 85179: gap of unknown length
 * 85180 88512: contig of 3333 bp in length
 * 88513 88612: gap of unknown length
 * 88613 92252: contig of 3640 bp in length
 * 92253 92352: gap of unknown length
 * 92353 94705: contig of 2353 bp in length
 * 94706 94805: gap of unknown length
 * 94806 98359: contig of 3554 bp in length
 * 98360 98459: gap of unknown length
 * 98460 103311: contig of 4852 bp in length
 * 103312 103411: gap of unknown length
 * 103412 107829: contig of 4418 bp in length
 * 107830 107929: gap of unknown length

Query Match 5.5%; Score 53; DB 2; Length 160634;
 Best Local Similarity 53.0%; Pred. No. 0.16;
 Matches 133; Conservative 0; Mismatches 117; Indels 1; Gaps 1;

```

QY      685  CCGCAGACCCGAGGAGGACCCCGGAGAGAGCCCACTGCGGAGGCGAGGCGGAAG 744
Db      35570 CCGCAGACCCGAGGAGGAGGACCCCGGAGAGAGCCCACTGCGGAGGCGAGGCGGAAG 35511
QY      745  ACAAGATACAGAGACCGGAGGACCAACGCGCGCCCGTGCATGAGACGAGAGACGCGGAGGGA 804
Db      35510 GAGGGGCGGCGCGCGCGGAGGAGGCGCGCGCGCGCGGAGACGAGCGCGCGGAGAGGC 35451
QY      805  GCGCAGACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 864
Db      35450 GCGCGGAGGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 35392
QY      865  CCGCAGCGCGCGCGCGCGCGCGCGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 924
Db      35391 CCGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 35332
QY      925  CCGACCGGCGCC 935
Db      35331 CCGCGGCGCGCG 35321

RESULT 6
AC090975/c 177485 bp DNA linear HTG 15-MAY-2002
LOCUS DEFINITION Mus musculus chromosome 17 clone RP23-290119 strain C57BL/6J,
AC090975 WORKING DRAFT SEQUENCE, 15 unordered pieces.
AC090975 AC090975.8 GI:19551112
VERSION HTG: HTGS PHASE1; HTGS DRAFT.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 177485)
AUTHORS Grills,G., Li,L., Montgomery,K.T., Chiu,D., Decker,J., Fusina,M.,
Goltz,J., Haider,A., Hall,L., Han,J., Ioshikhes,I.P., Lee,E.,
Perera,A., Shih,C., Thomas,E. and Kucheriapatti,R.
HIGH THROUGHPUT MOUSE SEQUENCING
Unpublished
2 (bases 1 to 177485)
JOURNAL Grills,G., Li,L., Montgomery,K.T., Chiu,D., Decker,J., Fusina,M.,
AUTHORS Goltz,J., Haider,A., Hall,L., Han,J., Ioshikhes,I.P., Lee,E.,
Perera,A., Shih,C., Thomas,E. and Kucheriapatti,R.
DIRECT SUBMISSION
Submitted (22-MAR-2001) Department of Molecular Genetics, Albert
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
Bronx, NY 10461, USA
3 (bases 1 to 177485)
JOURNAL Grills,G., Li,L., Montgomery,K.T., Chiu,D., Decker,J., Fusina,M.,
AUTHORS Goltz,J., Haider,A., Hall,L., Han,J., Ioshikhes,I.P., Lee,E.,
Perera,A., Shih,C., Thomas,E. and Kucheriapatti,R.
DIRECT SUBMISSION
Submitted (20-MAR-2002) Harvard Partners Center for Genetics and
Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA
02139, USA
COMMENT On Mar 20, 2002 this sequence version replaced gi:1748574.
-----Genome Center
Center: Harvard Partners Genome Center
Center Code: HPGC
Web site: http://www.hpcg.org/Sequence/mouse.html
Contact: hpcgcmendel.mgh.harvard.edu
-----Summary Statistics
Center project name: AEU
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye, 100#
*Consensus quality: 172552 at least Q20
*Consensus quality: 171794 at least Q30
*Consensus quality: 170494 at least Q40
*Estimated insert size: agarose-FP - N/A
**Estimated insert size: 177205 - sum-of-contigs
Quality coverage: agarose-FP - N/A
Quality coverage: 10.8 x in Q20 bases; sum-of-contigs estimation

```

```

FEATURES
Source
1. 177485 Mus musculus
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="17"
/clone="RP23-290119"
/sex="male"
1. 31079
/note="assembly_name:Contig161"
clone_end:SP6
vector_side:right"
31100_68493
/note="assembly_name:Contig160"
clone_end:IT7
vector_side:left"
68514_99930
/note="assembly_name:Contig159"
99951_130275
/note="assembly_name:Contig158"
130296_149687
/note="assembly_name:Contig157"
149708_159762
/note="assembly_name:Contig156"
159783_167610
/note="assembly_name:Contig155"
167631_173036
/note="assembly_name:Contig154"
173057_173359
/note="assembly_name:Contig153"
173380_174675
/note="assembly_name:Contig152"
174696_175576
/note="assembly_name:Contig151"

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```
misc_feature      .175722      175587="assembly_name:Contig150"  
misc_feature      .175743      176658 /note="assembly_name:Contig149"  
misc_feature      .176719      177353 /note="assembly_name:Contig146"  
misc_feature      .177404      177465 /note="assembly_name:Contig147"  
ORIGIN
```

Query Match	Best Local Similarity	5.5%;	Score 53;	DB 2;	Length 177485;
Best Local Similarity	49.8%;	Pred. No. 0.16;			
Matches 164;	Conservative 0;	Mismatches 160;	Indels 5;	Gaps 1;	
QY	628	AGTTGACCGCGCAGCGGCGCCGAAAGACAGGCTTTTCAGCGGATATGCCAAACACAGAGGCGG	687		
Db	68247	AGCCGGCCCGCCAGCGGAGACCGCGCCCGCGCGGGGGGGGCGGAGACCGACCGGGGGCGG	68188		
QY	688	CAGACCGGCGAGCGGACCCCGGGGGGAGAGACCCCACTGCGCGGACGGCGGAGGCGGAGAGACA	747		
Db	68187	CGCGGGCGCCCGCGCACCGGGAGAGGGGGGGGCGCGCGCCACCGCGGGGGGACACCGG	68128		
QY	748	GATACGAGGACCGCGAGGCCACACGCGCGCCCGTCATGACGAGACGCG-----CGAGG	802		
Db	68127	CGGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGGCGGAGGACGAGCGAGGGGCGGGACCGCGAGG	6806		
QY	803	GAGCGAGACCCCGAAGCGAGCGCGCCGCAAGACGACCGCGCAAGCGCACACGCGCCTAGAGCC	862		
Db	68067	GGGCGCCCGGCGCGGAGCGCGCGGGGGGGGGGCGGGGGGCGGCGCCACCCCGCGCG	68008		
QY	863	TGCGCCACCGCGCGGCCACGCGCGCGAGAGCGGGGCGGAGGACCGCAGAGACCGACCA	922		
Db	68007	CGCGGCGCCCGCGCGCGCCCGCCCGCACCGCGCCGCAACCCACCGCCACACCCACGACCC	67948		
QY	923	ACCAGCGGCGCCAGGCGGAGCGAGCGAC	951		
Db	67947	GCCCCCGCGAGCGCGCGCCCGACCGCCAC	67919		

RESULT	7
E32986	
LOCUS	E32986 10732 bp DNA linear PAT 18-JUN-2001
DEFINITION	Gene encoding cellulose synthetaser.
ACCESSION	E32986
VERSION	E32986.1 GI:13022340
KEYWORDS	JP 2000060568-A/1.
SOURCE	Vigna angularis (adzuki bean)
ORGANISM	Vigna angularis
	Burkholderia
	Viridiplantae; Streptophyta; Tracheophyta;
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
	Rosids; eurosoids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
	Vigna
REFERENCE	1 (bases 1 to 10732)
AUTHORS	Koichi,M.,Tomoniko,K., Shigeru,S. and Daisuke,S.
TITLE	Gene encoding cellulose synthetaser
JOURNAL	Patent: JP 2000060568-A 1 29-FEB-2000,
COMMENT	KOICHI MIZONO,MITSUO GYOUSAI SHOKUSETSU BIO KENKUSHO OS Vigna angularis

PF	26-AUG-1998 JP 1598239998
PR	
PI	KOICHI MIZUNO, TOMOHIKO KATO, SHIGERU SATO, DAISUKE SHIBATA PC
C12N15/09,A01H5/00,C07K16/40,C12N5/10,C12N9/10,C12P21/02// PC	
(C12N15/09,C12R1/91),(C12N5/10,C12R1/91),(C12P21/02,C12R1/91), PC	
C12N15/00,	
PC C12N5/00,(C12N15/00,C12R1/91),(C12N5/00,C12R1/91) CC	
FI	Key Location/Qualifiers
FT	CDS Location/Qualifiers (1). (3375).
FEATURES	
source	1..10732 /organism="Vigna angularis" /mol_type="genomic DNA"

ORIGIN	/db_xref="taxon:3914"
Query Match	5.3%; Score 50.8; DB 6; Length 10732;
Best Local Similarity	14.8%; Pred. No. 0.59;
Matches 135; Conservative 333; Mismatches 445; Indels 2; Gaps 2;	

Qy	41	GAATGATGCGTGGTCCGGCCCGAGGATGACGGGTCAAGCACTCTTGTAGCTTACGAGCCG	100
Db	9578	GAHSHTARGSRGTYSTYSGRSKCYSTSRVAGTAASRBSASMTTASASVAAWMTGYSAV	9637
Qy	101	ATGCAGTGTGAGATGCCACAGAACTACAGCCTGCACAGACTCTTAGCTGATGAAGG	160
Db	9638	AVASRVAGSYAAATHRGNSSRSPAASTRYSAVAATHGASNSRSRGTHRGTHYSGSRTR	9697
Qy	161	AGACACCGGGATGAAGACAGTGGTCCCTTACAGAAATCTCTATAGTGTCTTACACAG	220
Db	9698	RGGGYTSGRVAPASHSRGNTHRTHRGTHRTHRGVAVAASMTGMTSPASHSRBSRY	9757
Qy	221	CTCAGCCCCCTACCAATGGCGAGACTGACCGCTCAGGAGATCCCAAGGAGGGGT	280
Db	9758	SSRYSASNTHRCYSSRASHYSGTTHRVAGYRGNAVVAAYSTRGNAHSVATHRH	9817
Qy	281	CTCTCCATCCACACACACAGGCGAGTCAAAGCCCTTATATAGCCGATGCAACGC	340
Db	9818	SYGGYRASHSHSYVAALAAKHSVAAASMAHSASHSASRSRARGYSRVASASH	9877
Qy	341	ATGTAAAGCGCCTTACAAATAGATATTTGTGACTTGTGTAATGTCTTACATACAT	400
Db	9878	RGNASHRRGNCTYSHGYSR CYSSRSRVRVAVAASNNRPAAGYASVAAASRYSAG	9937
Qy	401	CAAAAGACATCTCACTACACAGAGTGCATGCCCTCGTGGCGGCGAATGCC	460
Db	9938	GRSRGNAGS-SRYSRGYTGVAGAGRAVAASTYRASSRSRBSAGNTSGYVAGASH	9996
Qy	461	ACTTCGCTCCGCGCTCCATATATATACATCATAGTTACACACCGGTGTCCGCGACAN	520
Db	9997	SRTRYASASGMYTTRHRTGYTRMTVAAGVAGHGNTRYTHHSRGRYSVAGHSHVAC	10056
Qy	521	TATGATGATATTTGCCCCGGGATCCACAGACATGATTTCCCCCGTGGCTGATCGAAT	580
Db	10057	STRASSRGTYSSTRASTYRGVASNMAAYTHTHTHAASTMTBASNHCYSHSRGNAY	10116
Qy	581	ATCGGGATCGCGGCATCAATTCGCGCGGAGCTACAGACTTAAATAATGACCGCGCA	640
Db	10117	ASHSRMTARASAAIYISGNHASTRASGTHRRAGSASGYGASNTGHHSHVAGISGARG	10176
Qy	641	GGCGCGGAGAACAGGCTTTGACGCGAATGCCAACAACAAGGGCCGACACCGCGAGC	700
Db	10177	HSRARGASNSRSRRGNVATHRTHRSR CYSHSAGNASGNSNGYGYRTYRANNGYSGHTH	10236
Qy	701	GACCGCGGGGAGAGAGCCCACTGCGGAGGGCGAGGAAGAGACAGTACGAGGACG	760
Db	10237	RATYSASYSRRAAAGSNGNTSIVAGARGYSGNSRPAATHRASAAAGSBMTASNNGNSRBY	10296
Qy	761	GAGCCACACCGCGCCCGTCATGACAGCGAGACGGCCGAGGAGCGCAGACCCGAAAGCA	820
Db	10297	STRASSRARGGNSHRYSGSPASNGYAGASNGYASASNGNTSASNSRASNAGHSTH	10356
Qy	821	GGCGCGCAAGAGGACCGGCGAAGCCACACGGGCGCTAGGCGCTGGG-CCAACGGCGGCA	879
Db	10357	RGAAGTSALAAHSASNTYVAHSFGGGASYSASNSASCTGGGAAYCTGGGNGNSR	10416
Qy	880	CGCGCGAGAGCGGGCGGAGACCGCAGAGAGACCGACCGAACCGGCCGAGGG	939
Db	10417	MTSRYSGCSYSSRSHASASYSASYSRRAGASVASTRGTHRAAASRGSAGCYSGNCT	10476
Qy	940	CACGACAGCCACCGCA 954	
Db	10477	HRASNGYSGNTYSA 10491	

RESULT 8

AF429315 125020 bp DNA linear PRI 18-JAN-2002
 LOCUS Home sapiens junctophilin 3 (JPH3) gene, partial cds.
 DEFINITION AF429315
 ACCESSION AF429315
 VERSION AF429315.1 GI:17646244
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 125020)
 Holmes, S.E., O'Hearn, E., Rosenblatt, A., Callahan, C., Hwang, H.S., Ingersoll-Ashworth, R.G., Fleisher, A., Stevanin, G., Brice, A., Potter, N.T., Ross, C.A. and Margolis, R.L.
 A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2
 Nat. Genet. 29 (4), 377-378 (2001)
 JOURNAL
 MEDLINE 21563737
 PUBMED 11694876
 REFERENCE 2 (bases 1 to 125020)
 Holmes, S.E., Ingersoll-Ashworth, R.G., Ross, C.A. and Margolis, R.L.
 AUTHOR TITLE Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
 JOURNAL
 FEATURES
 source 1..125020
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /map="16q24.3; between D16S520 and WI-12410"
 /note="Isolated from a patient with Huntington's Disease-like 2 (HDL2)"
 repeat_region complement(35581..35746)
 /rpt_type="tandem"
 /rpt_unit="ctg"
 complement(35507..35687)
 /gene="JPH3"
 /note="synonym: JF3"
 complement(35507..35687)
 /gene="JPH3"
 /product="junctophilin 3"
 complement(35507..35687)
 /protein_id="AA140941.1"
 /db_xref="GI:17646245"
 /translation="MSSGGRFMPDDGSGYCGWEDKRAHGHCCTGPKGGEGTGSWS HGFVTVGYTTPSNTYQGTWAGQKRGHIGLESKKWYVKGEMTHGFGRGVRECAAG NGAKYEGTWSNGLDGYETETSDG"
 ORIGIN
 Query Match 5.3%; Score 50.8; DB 9; Length 125020;
 Best local Similarity 12.0%; Pired. No. 0.51;
 Matches 54; Conservative 200; Mismatches 192; Indels 4; Gaps 1;
 498 CACACGGGTGTGGCGGACACANTATGATGATATTCGCCCGGATACAGACATGGA 557
 17136 SRSAMCCSYACKMCVSCYGMKCYVAAACRYGMSRSKRCMRGTYSTSGCCCT 17195
 558 TTTCCCGCGTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 617
 17196 TTTCCCGCGTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17255
 618 GACCTAAAGAGTTGACCGCGGACGCGGAGAGACAGGCTTTGACGAGATGCCAACA 677
 17256 GMSCTKXSSMTSMASYSYCMWSMYSCWBSMASRMSWSYVKKMKSSMRMSCTWKCCW 17315
 678 CAGAGGGCGGACAGCGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 737

Db 17316 CMKRCYCMRSMRSMWSYMYASWKSBSRGYCTRCYCMSSKSCYKSYMMRSMKNC 17375
 QY 738 GGAAGAGAGATPACAGAGAGCGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 797
 Db 17376 GMRKSRKSMWSMGASRSRSCYKCYKSMWSMWSMWSMWSMWSMWSMWSMWSMWSM 17435
 QY 798 CGAGGAGCGGACAGCGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 857
 Db 17436 RMRGKSMRSMRSMRSMRSMRSMRSMRSMRSMRSMRSMRSMRSMRSMRSMRSM 17491
 QY 858 GCGCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 917
 Db 17492 KCMSTYCMSTYCMSTYCMSTYCMSTYCMSTYCMSTYCMSTYCMSTYCMSTY 17551
 QY 918 CAGGAGCGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 947
 Db 17552 CTSRGAMSCMRRCYWRGASMSRAGSMR 17581
 RESULT 9
 AC142034/c 123570 bp DNA linear HTG 21-MAR-2003
 LOCUS Rattus norvegicus clone CH230-253B2, *** SEQUENCING IN PROGRESS
 DEFINITION *** 54 unordered pieces.
 ACCESSION AC142034
 VERSION AC142034.1 GI:29135505
 KEYWORDS HTG; HTGS_PHASE1.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 123570)
 Wuzny, D., Mathe, M., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amth, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, I., Bacc, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Crete, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Deederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Davies, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flegg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Gamba, R., Garcia, A., Garner, T., Garza, M., Gdresgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guetara, M., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, E., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Ledow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, M., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Louie, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Mallory, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M., McNeill, T., Meenan, E., Milosavljevic, A., Miner, G., Minje, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundada, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Naochelen, O., Okunolu, G., Olariunpase, A., Pal, S., Parke, K., Pastermak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindeexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzos, C., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reich, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodery, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savery, G., Scherer, S., Scott, G., Shateman, S., Shen, H., Shetty, J., Sivartsbeyn, A., Sison, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Unpublished
2 (bases 1 to 123570)
Morley, K.C.
Submitted (21-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, Ore
Baylor Plaza, Houston, TX 77030, USA
Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GQOT
Center clone name: CH230-253B2
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 91744 bases at least Q40
Consensus quality: 100630 bases at least Q30
Consensus quality: 105715 bases at least Q20
Estimated insert size: 90869; sum-of-coverage estimation
Quality coverage: 1x in Q20 bases; sum-of-coverage estimation

NOTE: Bactinated insert size may differ from sequence length
(see <http://www.hgsc.bcm.tmc.edu/docs/genbankdraftdata.html>).
NOTE: This is a 'working draft' sequence. It currently
consists of 54 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1419: contig of 1419 bp in length
1420 1519: gap of unknown length
1520 2583: contig of 1064 bp in length
2584 2684: gap of unknown length
2684 4098: contig of 1415 bp in length
4099 4198: gap of unknown length
4199 5410: contig of 1212 bp in length
5411 5510: gap of unknown length
5511 6856: contig of 1346 bp in length
6857 6956: gap of unknown length
6957 8164: contig of 1208 bp in length
8165 8264: gap of unknown length
8265 9302: contig of 1038 bp in length
9303 9402: gap of unknown length
9403 10679: contig of 1277 bp in length
10680 10779: gap of unknown length
10780 11779: contig of 1000 bp in length
11780 11879: gap of unknown length
11880 13188: contig of 1309 bp in length
13189 13288: gap of unknown length
13289 14534: contig of 1246 bp in length
14535 14634: gap of unknown length
14635 15660: contig of 1026 bp in length
15661 15760: gap of unknown length
15761 18008: contig of 2248 bp in length
18009 18108: gap of unknown length
18109 19731: contig of 1623 bp in length
19732 19832: gap of unknown length
19832 21073: contig of 1242 bp in length

21074 21173: gap of unknown length
21174 21176: contig of 1002 bp in length
21176 22275: gap of unknown length
22275 23779: contig of 1504 bp in length
23780 23879: gap of unknown length
23880 25886: contig of 2007 bp in length
25887 25986: gap of unknown length
25987 27663: contig of 1777 bp in length
27664 27863: gap of unknown length
27864 29028: contig of 1165 bp in length
29029 29128: gap of unknown length
29129 30386: contig of 1258 bp in length
30387 30486: gap of unknown length
30487 32341: contig of 1555 bp in length
32342 32441: gap of unknown length
32442 34065: contig of 1624 bp in length
34066 34165: gap of unknown length
34166 35979: contig of 1814 bp in length
35980 36079: gap of unknown length
36080 37895: contig of 1816 bp in length
37896 37995: gap of unknown length
37996 40144: contig of 2149 bp in length
40145 40244: gap of unknown length
40245 42243: contig of 1995 bp in length
42244 42343: gap of unknown length
42344 43405: contig of 1062 bp in length
43406 43505: gap of unknown length
43506 45570: contig of 2065 bp in length
45571 45670: gap of unknown length
45671 47982: contig of 2312 bp in length
47983 48082: gap of unknown length
48083 49373: contig of 1231 bp in length
49374 49473: gap of unknown length
49474 51351: contig of 1878 bp in length
51352 51451: gap of unknown length
51451 53148: contig of 1637 bp in length
53149 53248: gap of unknown length
53249 54816: contig of 1568 bp in length
54817 54916: gap of unknown length
54917 57844: contig of 2928 bp in length
57845 57944: gap of unknown length
57945 60066: contig of 2122 bp in length
60067 60166: gap of unknown length
60167 61534: contig of 1368 bp in length
61535 64603: gap of unknown length
64604 64503: gap of unknown length
64504 67778: contig of 3275 bp in length
67779 67878: gap of unknown length
67879 69295: contig of 1417 bp in length
69296 69395: gap of unknown length
69396 71592: contig of 2197 bp in length
71593 71692: gap of unknown length
71693 74596: contig of 2904 bp in length
74597 74696: gap of unknown length
74697 76912: contig of 2216 bp in length
76913 77012: gap of unknown length
77013 79045: contig of 2033 bp in length
79046 79145: gap of unknown length
79146 82134: contig of 2989 bp in length
82135 82234: gap of unknown length
82235 85286: contig of 3052 bp in length
85287 85386: gap of unknown length
85387 88447: gap of unknown length
88448 92453: contig of 4006 bp in length
92454 92553: gap of unknown length
92554 96685: contig of 4132 bp in length
96686 101131: gap of unknown length
101132 101231: contig of 4346 bp in length
101232 106183: gap of unknown length
106184 106283: gap of unknown length


```

Db      50718 ABDAGVCRCEWYKACADBCYKRDVSCWSAHVWYVYBBSKSSRBGRHMEACHKSM 50659
Qy      908 GGAGCGGACGACGAGCCGACCGGCC 934
Db      50658 CKSVYKVMWYKMKYCWKSCSWBCS 50632

RESULT 11
AD000671/c
LOCUS   AD000671
DEFINITION Homo sapiens DNA from chromosome 19-cosmid f24109 containing HRX2, genomic sequence.
ACCESSION AD000671
VERSION   AD000671.1 GI:1905893
KEYWORDS  chromosome 19.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 46251)
AUTHORS   Lamerdin,J.E.
TITLE     Direct Submission
JOURNAL   Submitted (27-NOV-1996) J.E. Lamerdin, Human Genome Center, Lawrence Livermore National Laboratory, 7000 East Ave, Livermore, CA, USA, 94551 jane@cgf.llnl.gov ow@cornak.llnl.gov
REFERENCE 2 (bases 1 to 46251)
AUTHORS   Lamerdin,J.E., McCreedy,P.M., Adamson,A.W., Burkhardt-Schultz,K., Garcia,E., Kyle,A., Ramirez,M., Stiiwagen,S., Ganes,J., Dargan,L., Bruce,R., Quan,G., Montgomery,M., Ow,D., Kobayashi,A., Olsen,A.O. and Carrico,A.V.
TITLE     Submitted (27-NOV-1996) Sequence analysis of a 1 Mb region in human 19q13.1
JOURNAL   GSDR:S:1010603.
COMMENT   constructed at LLNL from flow-sorted chromosomes from hybrid UV5HL9-5B, which carries chromosome 19 as its only human chromosome
           Overlaps cosmid R28052 (by 9.5 kb) to the left and is 1.4 kb to the left of cosmid F25965 map19q13.1 between D19S208 and CAMPS
           Human Genome Center
           Biology and Biotechnology Research Program
           Lawrence Livermore National Laboratory
           7000 East Avenue
           Livermore, CA 94550 USA.
FEATURES             source
     source           1..46251
                        Location/Qualifiers
                        1..46251
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="19"
                        /cell_line="UV5HL9-5B"
                        /cell_type="fibroblast"
                        /clone_id="LI19NC02 F2 chromosome 19-specific cosmid library"
     repeat_region    536..850
                        /note="Repeat match = HSA104796; putative"
                        /rpt_family="Alu"
     repeat_region    1142..1189
                        /note="Repeat match = HSA113971; putative"
                        /rpt_family="Alu"
     repeat_region    1153..1752
                        /note="Repeat match = HSA102794; putative"
                        /rpt_family="Alu"
     repeat_region    1484..2226
                        /note="Repeat match = HSA106431; putative"
                        /rpt_family="Alu"
     repeat_region    2108..2257
                        /note="Repeat match = HSA102931; putative"
                        /rpt_family="Alu"
     repeat_region    2561..2609
                        /note="Repeat match = HSA118610; putative"

/rpt_family="Alu"
2571..2730
/note="Repeat match = HSA106183; putative"
/rpt_family="Alu"
complement(2878..3169)
/note="Repeat match = HSA100369; putative"
/rpt_family="Alu"
complement(3131..3179)
/note="Repeat match = HSA105507; putative"
/rpt_family="Alu"
3221..3500.
/note="Repeat match = Alu; putative"
/rpt_family="Alu"
complement(4313..4415)
/note="Repeat match = HSA100485; putative"
/rpt_family="Alu"
complement(4334..4629)
/note="Repeat match = HSA106577; putative"
/rpt_family="Alu"
5110..5729
/note="Repeat match = HSA100284; putative"
/rpt_family="Alu"
6002..6035
/note="Repeat match = HSA117280; putative"
/rpt_family="Alu"
6011..6133
/note="Repeat match = HSA118162; putative"
/rpt_family="Alu"
6038..6271
/note="Repeat match = Alu; putative"
/rpt_family="Alu"
6790..7065
/note="ss region; putative"
7280..7543
/note="Repeat match = Alu; putative"
/rpt_family="Alu"
7475..7563
/note="Repeat match = HSA102791; putative"
/rpt_family="Alu"
7695..7936
/note="Repeat match = HSA102508; putative"
/rpt_family="Alu"
7847..7960
/note="Repeat match = HSA107958; putative"
/rpt_family="Alu"
10138..11022
/note="predicted exon, gra112exons human_1.3; frame=0, forward strand, quality=excellent; putative"
11736..11852
/note="similarity: sp|Q05516|PLZF_HUMAN; ZINC FINGER PROTEIN PLZF (PROMYELOCYTIC LEUKEMIA ZINC...); similarity: gi|913049; mPLZF(B)=promyelocytic leukemia zinc finger protein; putative"
11988..12077
/note="similarity: gi|545177; PLZF=Zinc finger protein(retinoic acid receptor alpha, RAR alpha...); similarity: sp|Q05516|PLZF_HUMAN; ZINC FINGER PROTEIN PLZF (PROMYELOCYTIC LEUKEMIA ZINC...); similarity: gi|913049; mPLZF(B)=promyelocytic leukemia zinc finger protein; putative"
12608..12748
/note="predicted exon, gra112exons human_1.3; frame=1, forward strand, quality=good; putative"
13530..13892
/note="predicted exon, gra112exons human_1.3; frame=2, forward strand, quality=good; putative"
15698..15771
/note="predicted exon, gra112exons human_1.3; frame=0, forward strand, quality=good; putative"
16358..16532
/note="predicted exon, gra112exons human_1.3; frame=2, forward strand, quality=good; putative"
16422..16502

```


		/note="similarity: p t B48205; A11-1 protein -GTE form - mouse (fragment); similarity: D14540; (D14540) MLL [Homo sapiens]; similarity: L22179; (L22179) MLL-AF4 der(11) fusion protein [Homo sapiens]; putative"
		16656. .16703
	misc_feature	/note="similarity: D14540; (D14540) MLL [Homo sapiens]; similarity: L22179; (L22179) MLL-AF4 der(11) fusion protein [Homo sapiens]; similarity: p t B48205; A11-1 protein -GTE form - mouse (fragment); putative"
		16806. .16865
	misc_feature	/note="similarity: p t B48205; A11-1 protein -GTE form - mouse (fragment); similarity: L22179; (L22179) MLL-AF4 der(11) fusion protein [Homo sapiens]; similarity: D14540; (D14540) MLL [Homo sapiens]; putative"
		17224. .17315
	misc_feature	/note="predicted exon, graitlexons_human_1.3; frame=2, forward strand, quality=excellent; putative"
		17305. .17454
	misc_feature	/note="ss region; putative"
	repeat_region	/note="repeat match = HSAU06523; putative"
		17843. .17983
	misc_feature	/note="predicted exon, graitlexons_human_1.3; frame=1, forward strand, quality=excellent; putative"
		17998. .18065
	misc_feature	/note="ss region; putative"
		18079. .18228
	misc_feature	/note="predicted exon, graitlexons_human_1.3; frame=0, forward strand, quality=excellent; putative"
		18088. .18264
	misc_feature	/note="similarity: D14540; (D14540) MLL [Homo sapiens]; similarity: p t A44265; trithorax-homologous protein (translocation) HRX - human; similarity: L04284; (L04284) HRX [Homo sapiens]; putative"
		18088. .18222
	misc_feature	/note="similarity: L22179; (L22179) MLL-AF4 der(11) fusion protein [Homo sapiens]; putative"
		18661. .18786
	misc_feature	/note="similarity: p t A44265; trithorax-homologous protein (translocation) HRX - human; similarity: L22179; (L22179) MLL-AF4 der(11) fusion protein [Homo sapiens]; similarity: L04284; (L04284) HRX [Homo sapiens]; similarity: D14540; (D14540) MLL [Homo sapiens]; putative"
		18664. .18785
	misc_feature	/note="predicted exon, graitlexons_human_1.3; frame=0, forward strand, quality=good; putative"
		18958. .19027
	misc_feature	/note="predicted exon, graitlexons_human_1.3; frame=1, forward strand, quality=excellent; putative"
		19371. .19517
	misc_feature	/note="predicted exon, graitlexons_human_1.3; frame=1, forward strand, quality=excellent; putative"
		19579. .19626
	misc_feature	/note="predicted exon, graitlexons_human_1.3; frame=0, forward strand, quality=good; putative"
		19727. .19775
	repeat_region	/note="repeat match = HSAL03439; putative"
		19739. .20004
	repeat_region	/rpt_family="Alu"
		/note="repeat match = HSAL01336; putative"
		/rpt_family="Alu"
		5.2%; Score 50.2; DB 9; Length 46251;
	Query Match	Best Local Similarity 4.7%; Pred. No. 0.74;
	Matches 148; Conservative 0; Mismatches 163; Indels 0; Gaps 0;	
632	GACCGCCAGCGGCCGAAGAACAAGCTTTGCAGCAGATGCCAAACACAGAGCCGACAGA	691
Db	13619 GCCCGCCGCGCGCCCGCGCGCGCGGGAGAACGCGCCCGCGAGACGAGCCCGAG	13560
692	C CGGAGGCAACCCCGGCGGAGAGCCCACTTCGCGCAGGCGGAGGACGAAGACAGATA	751

Db	13559	GCTACTGTCGCGCCCGCCGCGCCGCGCCATCTTTGGCACCCTGATGAGAGGCGCCGGGAGAGCG	135000
Qy	752	CGAGACGCGAGCCACACGCGCGCCCGTCATGAGACGAGACGCGCCGAGGAGCGGCAGACA	811
Db	13449	GGGGAGAGGGCGCGCCGTGTGCGAGGGCCGGGGGAGAGGAGAGAGGGGGGCGGGCGGG	13440
Qy	812	CCCGAAGACAGCGCGCCGAAGAGCCACCGGCGGCAGACACAGCGCCCTTAGCGCCGACAG	871
Db	13439	GCGGCGCGCGCGCAGACACACACAGCGCCCGCGCGGAGACCGGGGCGCGGCGCGGC	13380
Qy	872	GCGGCGCACCGCGCAGAGCGCGGCGCGCAGACCGCAGACCGACACGACCGACCGC	931
Db	13379	ACGGCGCAGGGCACACCGAGGAGAGGAGTCGCGGCGCGCGCGCGCCCTCGACCGGACCT	13320
Qy	932	GCCCCAGGCGAG 942	
Db	13319	GGGGAGGACAG 13309	
RESULT 12			
LOCUS	AC022315		
DEFINITION	AC022315	194835 bp	DN
	AC022315	194835 bp	linear
	AC022315	194835 bp	HTG 25-APR-2002
	AC022315	194835 bp	WORKING DRAFT SEQUENCE,
	AC022315	194835 bp	18 unordered pieces.
ACCESSION	AC022315		
VERSION	AC022315.45	GI:19795245	
KEYWORDS	HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 194835)		
AUTHORS	Adola,A.P., Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D.,		
	Federpspiel,N., Glukhov,S., Hansen,N., Herman,Z.S., Hyman,R.,		
	Mao,J., Komp,C., Kotlier,S., Lam,B., Marathe,R., Miranda,M.,		
	Morehouse,A.J., Nguyen,M., Oefner,P., Palm,C.J., Ramirez,D.,		
	Southwick,A.W., Webb,C., Wilhelmy,J., Yu,S. and Davis,R.W.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 194835)		
AUTHORS	Bruno,D., Conn,L., Dela Rosa,M., Federpspiel,N., Foreman,P.,		
	Glukhov,S., Hansen,N., Hyman,R., Mao,J., Marathe,R.,		
	Morehouse,A.J., Oefner,P., Palm,C.J., Ramirez,D., Wilhelmy,J.,		
	Yu,S. and Davis,R.W.		
TITLE	Direct Submision		
JOURNAL	Submitted (30-JAN-2000) DNA Sequencing and Technology Center,		
	Stanford University, 855 California Avenue, Palo Alto, CA 94304,		
	USA		
REFERENCE	3 (bases 1 to 194835)		
AUTHORS	Bruno,D., Conn,L., Dela Rosa,M., Federpspiel,N., Foreman,P.,		
	Glukhov,S., Hansen,N., Hyman,R., Mao,J., Marathe,R.,		
	Morehouse,A.J., Oefner,P., Palm,C.J., Ramirez,D., Wilhelmy,J.,		
	Yu,S. and Davis,R.W.		
TITLE	Direct Submision		
JOURNAL	Submitted (25-APR-2002) DNA Sequencing and Technology Center,		
	Stanford University, 855 California Avenue, Palo Alto, CA 94304,		
	USA		
COMMENT	On Dec 21, 2001 this sequence version replaced gi:17063116.		
	----- Genome Center		
	Center: Stanford DNA Sequencing and Technology Development		
	Center		
	Center code: SDS/RDC		
	Web site: http://sequence-www.stanford.edu/group/human/		
	Contact: hum-info@sequence.stanford.edu		
	----- Project Information		
	Center project name: RP11-38C1		
	Center Clone name: RP11-38C1		
	----- Summary Statistics		
	Sequencing Vector: M13mp18; X02513; 96% of reads		
	Sequencing Vector: plasmid; plasmid accession; 4% of reads		
	Chemistry: Dye-terminator; 12% of reads		
	Chemistry: Dye-terminator Big Dye; 87% of reads		
	Assembly program: Phrap; version 0.990319		

Consensus quality: 184441 bases at least Q40
 Consensus quality: 186605 bases at least Q30
 Consensus quality: 187666 bases at least Q20
 Insert size: 168558; agarose-fp
 Insert size: 193135; sum-of-contigs
 Quality coverage: 14.4x in Q20 bases; agarose-fp
 Quality coverage: 12.6x in Q20 bases; sum-of-contigs.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 18 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1604: contig of 1604 bp in length
 * 1605 1700: gap of unknown length
 * 1705 3770: contig of 2066 bp in length
 * 3771 3870: gap of unknown length
 * 3871 5909: contig of 2039 bp in length
 * 5910 6009: gap of unknown length
 * 6010 8095: contig of 2086 bp in length
 * 8096 8195: gap of unknown length
 * 8196 10705: contig of 2510 bp in length
 * 10706 10805: gap of unknown length
 * 10806 12952: contig of 2147 bp in length
 * 12953 13052: gap of unknown length
 * 13053 16988: contig of 3936 bp in length
 * 16989 17088: gap of unknown length
 * 17089 20048: contig of 2860 bp in length
 * 20049 20149: gap of unknown length
 * 20149 23005: contig of 2857 bp in length
 * 23006 23105: gap of unknown length
 * 23106 27314: contig of 4209 bp in length
 * 27315 27415: gap of unknown length
 * 27415 31443: contig of 4029 bp in length
 * 31443 31543: gap of unknown length
 * 31544 37789: contig of 6246 bp in length
 * 37790 37890: gap of unknown length
 * 37890 44274: contig of 6385 bp in length
 * 44275 44374: gap of unknown length
 * 44375 53077: contig of 8703 bp in length
 * 53078 53177: gap of unknown length
 * 53178 60886: contig of 7709 bp in length
 * 60887 60986: gap of unknown length
 * 60987 77446: contig of 16460 bp in length
 * 77447 77546: gap of unknown length
 * 77547 107542: contig of 29996 bp in length
 * 107543 107643: gap of unknown length
 * 107643 194835: contig of 87193 bp in length.

FEATURES

source
 1. 194835
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="19"
 /clone="RP11-38C1"
 /clone_1lb="RP11 human BAC library 11"
 1. 1604
 /note="assembly_name:Contig33"
 misc_feature
 1705..3770
 /note="assembly_name:Contig44"
 misc_feature
 3871..5909
 /note="assembly_name:Contig48"
 misc_feature
 6010..8095
 /note="assembly_name:Contig50
 clone_end:77"
 8196..10705
 /note="assembly_name:Contig51"
 misc_feature
 10806..12952
 /note="assembly_name:Contig52"
 misc_feature
 13053..16988
 /note="assembly_name:Contig53"

misc_feature 17089..20048
 /note="assembly_name:Contig54"
 misc_feature 20149..23005
 /note="assembly_name:Contig55"
 misc_feature 23106..27314
 /note="assembly_name:Contig56"
 misc_feature 27415..31443
 /note="assembly_name:Contig57"
 misc_feature 31544..37789
 /note="assembly_name:Contig58"
 misc_feature 37890..44274
 /note="assembly_name:Contig59"
 misc_feature 44375..53077
 /note="assembly_name:Contig60"
 misc_feature 53178..60886
 /note="assembly_name:Contig61"
 misc_feature 60987..77446
 /note="assembly_name:Contig62"
 misc_feature 77547..107542
 /note="assembly_name:Contig63
 clone_end:SP6"
 107643..194835
 /note="assembly_name:Contig64"

ORIGIN

Query Match 5.2%; Score 50.2; DB 2; Length 194835;
 Best Local Similarity 47.6%; Pred. No. 0.67;
 Matches 148; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 632 GACCGCGCAGCGCGCGCAAGAACAGGCTTTCAGCGCAATGCCAAACAGAGGCGCGCA 691
 DB 184206 GCGCCCG 184265
 QY 692 CCGGACG 751
 DB 184266 GCAACTGCG 184325
 QY 752 CGAGACG 811
 DB 184326 GGGGGAGGGCG 184385
 QY 812 CCGGACG 871
 DB 184386 GCG 184445
 QY 872 GCG 931
 DB 184446 AGCGCGAGGCGCACACGCGGAGGAGGAGTCGCGCGCGCGCGCGCGCGCGCGCG 184505
 QY 932 GCCCAGGCGAG 942
 DB 184506 GGGGAGGCGAG 184516

RESULT 13

AC143344/C 152686 bp DNA linear HTG 10-APR-2003
 LOCUS AC143344
 DEFINITION Macaca mulatta clone CH250-268P23, *** SEQUENCING IN PROGRESS ***.
 ACCESSION AC143344
 VERSION AC143344.1 GI:29611668
 KEYWORDS HTG; HTGS PHASE2; HTGS_PG1.
 SOURCE Macaca mulatta (rhesus monkey)
 ORGANISM Macaca mulatta
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 Cercopithecinae; Macaca.
 1 (bases 1 to 152686)
 Ceuro/M. and Milosavljevic/A.
 Pooled genomic indexing (Pgi): mathematical analysis and experiment
 design
 (in) Guigo R. and Gusfield D. (Eds.):
 ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI
 2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-26;

REFERENCE

ATTIORS
 TITLE
 JOURNAL

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 14:04:56 ; Search time 364.27 Seconds
(without alignments)
1160.755 Million cell updates/sec

Title: US-10-076-747-7
Perfect score: 957
Sequence: 1 ttggcgcttaactatgcatg.....ggcagcagagcaccgcagct 957

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373663 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_292an04:*
1: geneseq1360s:*
2: geneseq1390s:*
3: geneseq2000s:*
4: geneseq2001as:*
5: geneseq2002s:*
6: geneseq2003as:*
7: geneseq2003bs:*
8: geneseq2003cs:*
9: geneseq2003ds:*
10: geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	956	99.9	957	ABX92193	ABX92193 Human ova
2	809.6	84.6	1460	ABX92194	ABX92194 Human ova
3	68.2	7.1	2067	ABA09659	ABA09659 Human bon
4	55.2	5.8	349	ABT07668	ABT07668 Human bre
5	55	5.7	1190	AA883947	AA883947 DNA encod
6	51	5.3	908	AA887524	AA887524 DNA encod
7	50.8	5.3	10732	AAA10594	AAA10594 Gene enco
8	49.6	5.2	2000	ADA71938	ADA71938 Rice gene
9	48.4	5.1	114955	AA53491	AA53491 Human ade
10	44.4	4.6	1451	AAA02527	AAA02527 Human col
11	44.2	4.6	87980	AA53223	AA53223 (5 of
12	43	4.5	366	ABX43944	ABX43944 Bovine ES
13	43	4.5	764	ADA10928	ADA10928 Human CDN
14	43	4.5	6225	AA55273	AA55273 Human enz
15	43	4.5	6225	AA54721	AA54721 Human enz
16	43	4.5	6225	AA54721	AA54721 Human mul
17	43	4.5	6225	AB296537	AB296537 Human nuc
18	43	4.5	35459	AB297130	AB297130 Human nuc
19	42.8	4.5	5760	AB578661	AB578661 M. echino
20	42.4	4.4	2919	ACA26704	ACA26704 Prokaryot
21	42.4	4.4	5059	AA584332	AA584332 Steath v
22	42.4	4.4	5617	AA571163	AA571163 Partial h
23	42.4	4.4	8091	AA57001	AA57001 Human Not

C 24	42.4	4.4	8091	6	AB234974	AB234974 Human gen
C 25	42.4	4.4	8257	6	AB570408	AB570408 Human bon
C 26	42.4	4.4	168174	6	ABT11173	ABT11173 Human S-1
C 27	42.4	4.4	168273	6	ABT11114	ABT11114 Human S-1
C 28	42.2	4.4	895	6	ABV83635	ABV83635 Human bre
C 29	42.2	4.4	4705	7	ACA61566	ACA61566 Streptomy
C 30	42.2	4.4	35384	4	AA521436	AA521436 Human enz
C 31	42.2	4.4	109519	5	AA506693	AA506693 Micromono
C 32	42	4.4	691	6	ABQ40803	ABQ40803 Oligonuc
C 33	42	4.4	691	6	ABQ40802	ABQ40802 Oligonuc
C 34	42	4.4	756	6	ABA99456	ABA99456 Actinopla
C 35	42	4.4	38064	6	ABA99469	ABA99469 Actinopla
C 36	41.8	4.4	712	6	ABQ40859	ABQ40859 Oligonuc
C 37	41.8	4.4	4488	2	ABQ40858	ABQ40858 Oligonuc
C 38	41.6	4.3	4488	2	AA555520	AA555520 Human thr
C 39	41.4	4.3	4570	9	ABE57372	ABE57372 Human gen
C 40	41.4	4.3	37314	4	AAK71358	AAK71358 Human imm
C 41	41.4	4.3	172570	6	ABQ88207	ABQ88207 Human ost
C 42	41.2	4.3	920	4	AAH21491	AAH21491 Barley al
C 43	41.2	4.3	1164	4	AA551496	AA551496 Pseudomon
C 44	41.2	4.3	1164	7	ACA19500	ACA19500 Prokaryot
C 45	41.2	4.3	154746	6	AA255519	AA255519 Human her

ALIGNMENTS

RESULT 1	ABX92193	ABX92193 standard; CDNA, 957 BP.
ID	ABX92193	
AC	ABX92193	
XX		
DT	08-MAY-2003	(first entry)
XX		
DE	Human ovarian specific nucleic acid DEX0310_7.	
XX		
XX		
KW	Human; gene; ss; ovarian specific nucleic acid; OSNA; ovarian cancer;	
KW	non-cancerous ovarian disease; gene therapy; vaccine; cyostatic;	
KW	gynaecological.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200292785-A2.	
PD		
XX		
XX		
PF	13-FEB-2002; 2002WO-US022271.	
XX		
PR	13-FEB-2001; 2001US-0268290P.	
PR	15-FEB-2001; 2001US-0268834P.	
XX		
XX	(DIAD-) DIADEXUS INC.	
PI	Saleda S, Macina RA, Hu P, Recipon H, Karra K, Caffery R;	
PI	Sun Y, Liu C;	
XX	WPI; 2003-120677/11.	
XX	P-PSDB; AB061022.	
PT	New isolated OSNA nucleic acid and encoded polypeptide, useful for	
PT	identifying, diagnosing, monitoring, staging, imaging and treating	
PT	ovarian cancer and non-cancerous diseases in ovarian tissues.	
PS	Claim 1, Page 145, 224pp; English.	
XX		
CC	The invention relates to a new isolated nucleic acid termed ovarian	
CC	specific nucleic acid (OSNA) comprising: (a) a nucleic acid molecule that	
CC	encodes any of 53 fully defined protein sequences appearing as AB061018-	
CC	AB061070 (termed ovarian specific proteins, OSP); (b) any of 76 fully	
CC	defined nucleotide sequences appearing as ABX92187-ABX92262; or (c) a	
CC	sequence having at least 60% sequence identity to the nucleic acid	
CC	molecule of (a) or (b). Also included are a method for determining the	


```

DB 634 AAGCTACAGCTGCCAGGTGAC-----GCATGAAGGAG-CACCGTGATGAAGACGT 685
QY 183 GGTCCCCCTACAGAAATGTTCTATAGTTCTCTMAACAGCTGAGCCCCACTACCAATGGC 242
DB 686 GGTCCCCCTACAGAAATGTTCTATAGTTCTCTMAACAGCTGAGCCCCACTACCAATGGC 745
QY 243 GAGACTAGCAGCTGCGAGGAGTCCCAAGGGGAGAGGGGTCTCTCATCCACACACCAAG 302
DB 746 GAGACTAGCAGCTGCGAGGAGTCCCAAGGGGAGAGGGGTCTCTCATCCACACACCAAG 805
QY 303 GCGAGTCAAAAGCCCTTATCATCGCGATGTCGACGTATGTAAGGCGCTACAAATTA 362
DB 806 GCGAGTCAAAAGCCCTTATCATCGCGATGTCGACGTATGTAAGGCGCTACAAATTA 865
QY 363 GATATTCGCACTGTTGTAATGCTTACTACTATAATAAAGACATATCTCACTACA 422
DB 866 GATATTCGCACTGTTGTAATGCTTACTACTATAATAAAGACATATCTCACTACA 925
QY 423 CGGAGTCGATGATCACCGGTCCGTGCGGGGGAATGCACTTGGCTGTCGTCGCAAT 482
DB 926 CGGAGTCGATGATCACCGGTCCGTGCGGGGGAATGCACTTGGCTGTCGTCGCAAT 985
QY 483 GATGACTATAGTTACACACCGGTGTGTGGCGGCACANCTATGATGATATTTGCCCGG 542
DB 986 GATGACTATAGTTACACACCGGTGTGTGGCGGCACANCTATGATGATATTTGCCCGG 1045
QY 543 ATCAAGACCATGATTTCCCGGTGATGATGATGATGATGATGATGATGATGATGAT 602
DB 1046 ATCAAGACCATGATTTCCCGGTGATGATGATGATGATGATGATGATGATGATGAT 1105
QY 603 CGCCCGGAGGTACAGACTTAAAGTTGAACCGCGGAGCGCGGAGGAGGAGGAGGAGG 662
DB 1106 CGCCCGGAGGTACAGACTTAAAGTTGAACCGCGGAGCGCGGAGGAGGAGGAGGAGG 1165
QY 663 ACGGAATGCCAACAACAGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 722
DB 1166 ACGGAATGCCAACAACAGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1225
QY 723 CTGCGGCGAGGCGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 782
DB 1226 CTGCGGCGAGGCGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1285
QY 783 TGAAGCGGAGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 842
DB 1286 TGAAGCGGAGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1345
QY 843 AGCCACAGCGGCGCTAGGCGCTGCGGCGAGCGGCGGAGGAGGAGGAGGAGGAGGAGG 902
DB 1346 AGCCACAGCGGCGCTAGGCGCTGCGGCGAGCGGCGGAGGAGGAGGAGGAGGAGGAGG 1405
QY 903 CCGCAGAGAGCGGAGCAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCT 957
DB 1406 CCGCAGAGAGCGGAGCAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCT 1460

RESULT 3
ABA09659/c
ID ABA09659 standard; DNA; 2067 BP.
XX
XX ABA09659;
AC
XX
XX 15-JAN-2002 (first entry)
DE
XX
XX Human bone marrow expressed oligonucleotide SEQ ID NO: 168.
XX
XX Human; bone marrow; cytosolic; antirheumatic; antiarthritic; vulnery;
KM antinflammatory; antibacterial; immunosuppressive; vasotropic; cancer;
KM antiParkinsonian; neuroprotective; nootropic; haemostatic; osteopathic;
KM antidiabetic; antidiabetic; antidiabetic; antidiabetic; antidiabetic;
KM immunostimulant; analgesic; cerebroprotective; antihaemic; infection;
KM nervous system disorder; autoimmune disorder; inflammation; allergy; ds.
XX

```

```

OS Homo sapiens.
XX
XX WO200174836-A1.
XX
XX 11-OCT-2001.
XX
XX
XX 30-MAR-2001; 2001WO-US010472.
XX
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX 30-NOV-2000; 2000US-02505839.
XX
XX (HYSE-) HYSEQ INC.
PA
PI Tang YT, Liu C, Dermanac RT, Ford JE, Boyle BJ;
XX
XX WPI; 2001-626375/72.
XX
XX
XX New bone marrow-expressed nucleic acids and polypeptides, useful for
XX diagnosis, treatment of inflammatory, autoimmune, neurological, cancer
XX PT and increasing hematopoiesis, stem cell survival and bone growth and
XX remodeling.
XX
XX Claim 1; Page 248-249; 380pp; English.
XX
XX
XX The present invention relates to bone marrow expressed polynucleotides
XX CC and proteins. These sequences can be used in the treatment of
XX CC inflammatory conditions (eg arthritis, Crohn's disease), cancer, central
XX CC and peripheral nervous system diseases and neuropathies, such as
XX CC Alzheimer's, Parkinson's and Huntington's diseases, spinal cord
XX CC disorders, head trauma, cerebrovascular diseases, myeloid and lymphoid
XX CC cell disorders, platelet disorders, stem cell disorders, bone
XX CC degenerative disorders, autoimmune disorders, for example multiple
XX CC sclerosis, diabetes and arthritis, viral and bacterial infections,
XX CC allergies and blood coagulation disorders. The present sequence is a DNA
XX
XX
XX Sequence 2067 BP; 418 A; 579 C; 620 G; 450 T; 0 U; 0 Other;
XX
XX
XX Query Match 7.1%; Score 68.2; DB 5; Length 2067;
XX Best Local Similarity 63.6%; Pred. No. 6.6e-08;
XX Matches 138; Conservative 0; Mismatches 73; Indels 6; Gaps 2;

QY 79 GCTATCTGAGCTGACGCGCTGATGAGTGTGAGTGTCCACAGAGCTACAGCTGAC 138
DB 987 GCTATCTGAGCTGAGCGCTGACGCGCTGAGTGTGAGTGTCCACAGAGCTACAGCTGAC 930
QY 139 AGAGTCTTAGCTGATGAGGAGGAGACCGGTGATGAGGAGGAGTGTCCCTACAGAAAT 198
DB 929 -CAGGTCACCGGCACTTGAAGGAGGAGCACCGTGGAGAGAGACAGGGGCCCCCTACAGCAAT 871
QY 199 GTTCTATAGTCTCTTAACAGGCTGAGCCGCCCACTACCAATGGCGAGACTAGACGCTGC 258
DB 870 GTTCTATAGTCTCTTAACAGGCTGAGCCGCCCACTACCAATGGCGAGACTAGACGCTGC 814
QY 259 AGGATCCCAAGGAGGAGGAGGCTCTCGATCCACAC 295
DB 813 GGAATCCCAAGGAGGAGGAGGCTCTCTCTCCACCC 777

RESULT 4
ABT07668
ID ABT07668 standard; cDNA; 349 BP.
XX
XX ABT07668;
AC
XX
XX 14-NOV-2002 (first entry)
DE
XX
XX Human breast cancer associated coding sequence SEQ ID NO: 147.
XX
XX Human; breast specific gene; breast specific protein; breast cancer;
KM gene therapy; cytosolic; gene; ss.
XX

```


PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 PS Claim 1, SEQ ID NO 15751, 103pp; English.
 XX
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

[illegible]

```

OY      259   AGGGATCCCAAGGAGAGGGGTCTTCCATCCACACC    295
          |||||
Db       846   AGATTGCCAGGGAGAGGGGTCTTCTCCTCCACCCC    882
          |||||

RESULT 6
AAS87524
ID      AAS87524 standard; cDNA; 908 BP.
XX
XX      AAS87524;
AC
XX
DT      13-FEB-2002 (first entry)
XX
XX      DNA encoding novel human diagnostic protein #23328.
DE
XX
XX      Human chromosome mapping; gene mapping; forensic;
KW      food supplement; medical imaging; diagnostic; genetic disorder; ss
XX
XX      Homo sapiens.
OS
XX      WO200175067-A2.
PN
PD      11-OCT-2001.
PP
30-MAR-2001; 2001WO-US008631.
```

XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX MPI: 2001-639362/73.
DR P-PSDB; ABG23337.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 23328; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 908 BP; 213 A; 290 C; 228 G; 177 T; 0 U; 0 Other;
Query Match 5.3%; Score 51; DB 5; Length 908;
Best Local Similarity 68.5%; Pred. No. 0.0019;
Matches 146; Conservative 0; Mismatches 55; Indels 12; Gaps 5;
QY 74 CAGCAGCTCTGAGCCTGACCGCTGATGAGTGAAGATCCACAGAGCTACAGCC 133
DB 623 CAGCAGCTCTGAGCCTGAGCCTGA-GCAGTGGGA-AGTCCACAGAGCTACAGCT 680
QY 134 TGCACAGAGCTTGAAGTGAAGAGACACCGTGAAGAGAGAGTGTCTCCCTTACA 193
DB 681 TGC---CAGGTCACAGGACATGAAGGGAGCACCGTGGAGAGACAGATGGGCCCTTACA 737
QY 194 GAATGTTCTATAGGTTCTCTACACGCTCAGCCGCCCTACACATGCGAGAGCTAGAC 253
DB 738 G-AAAGTTTCATAGGTTTCTCAACCCCTCACTCCCCCA-----CGGAAACTTAG 790
QY 254 GCTGCAGGATCCCAAGGAGAGGGGATCTCTCC 286
DB 791 ACTTCAGAAATCCACGGGAGGGGATTTCTCTCC 823
RESULT 7
AAA10594
ID AAA10594 standard; DNA; 10732 BP.
XX
XX AAA10594;
AC
XX
XX 29-JUN-2000 (first entry)
DT
XX
DE Gene encoding a subunit of cellulose synthase.

KW Cellulose synthase; cellulose production; increase yield; ds.
XX
XX Vigna angularis.
OS
XX JP2000060568-A.
PN
XX
XX 29-FEB-2000.
PD
XX
XX 26-AUG-1998; 98UP-00239998.
PF
XX
XX 26-AUG-1998; 98UP-00239998.
PR
XX
XX 26-AUG-1998; 98UP-00239998.
XX
XX (MIZU/) MIZUNO K.
PA (OJIP) OJI PAPER CO.
XX
XX MPI: 2000-342371/30.
DR P-PSDB; AAY85179.
XX
XX A gene encoding a cellulose synthetic equipment - for the improvement in
PT the amount of cellulose synthesized in a plant body.
PT
XX
XX Claim 2; Page 14-21; 32p; Japanese.
PS
XX
XX This sequence represents a gene encoding a subunit of the cellulose
CC synthase complex of Vigna angularis. The invention relates to subunits of
CC cellulose synthetic equipment, that can be used to increase the amount of
CC cellulose synthesised by a plant. The proteins and genes encoding them
CC can also be used to improve the properties of the cellulose being
CC produced by a plant
XX
SQ Sequence 10732 BP; 3149 A; 1212 C; 2074 G; 2046 T; 0 U; 2251 Other;
Query Match 5.3%; Score 50.8; DB 3; Length 10732;
Best Local Similarity 14.8%; Pred. No. 0.0047;
Matches 135; Conservative 333; Mismatches 445; Indels 2; Gaps 2;
QY 41 GATGATGCTGCTGCGCGCGGAGTACCGGCTGACGAGCTATGAGCTTGAACGCTG 100
DB 9578 GASHSHARGSGRGYSYSGRSRCSYSSRVAGYAASRAASAMNTASASVAGMTGSAV 9637
QY 101 ATGCAGTGTGAGAGTCCACAGAGCTACAGGCTGACAGAGTCTAGCTGATGAAGG 160
DB 9638 AVASRVAGTGAATHRGNSRRAASRYSAVATHRGASRSRGTHRYSGSRTTR 9697
QY 161 AGACACGCTGATGAAGACAGTGTCCCTACAGAAATGTTCTAATGTTCTTACACG 220
DB 9698 RGGGYSGRVARASHSSRGTHRTGTRHRGVAVAASMTGRMTSRASHSSRASRY 9757
QY 221 CTGAGCCCCACATCAATGCGAGACTGACACCGCTGACAGGATCCCAAGGAGAGG 280
DB 9758 SSRYASNTHRCTSSPASHYSGYTHRVANGYTRGNASVAVASTHGNASHSVATHR 9817
QY 281 CTCTCCATCCACACACAGAGGCGAGTCAAGCCCTTATCATCGCGATGTCAGCTC 340
DB 9818 SYSGYASASHASYSVAAAAGHSAASHASHASASRSRARGYSRVASHASTH 9877
QY 341 ATGTAAGAAGGCTCAAAATAGATATTCTGCACTGTGTAATCTCTCATACATATA 400
DB 9878 RGNASNRNGCTSHSYSSRYSRRAASRVAAVNASNRRAAGASVAAASSRASAG 9937
QY 401 CAAGAACAATCACTACACAGGATGATGATGATGATGATGATGATGATGATGATGAT 460
DB 9938 GRSRNGAS-SRYSRSGYSGVAGAGRAVAAVASTYRASRSRAGNSGAGVAGASH 9996
QY 461 ACTTCGCTCTGCTGCTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 520
DB 9997 SRTYRASNSGVTYTRTHRYTMTVAAGVAGVAGVAGVAGVAGVAGVAGVAGVAG 10056
QY 521 TATGAGTGAATATGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 580
DB 10057 STYRASRGYSTRASRYTRGASNAAGYTRTHRAAASMTGTAASNHGCTSHSRNABV 10116
QY 581 ATGGGATGACGCGGATCAAAATTCGCGGAGCTACAGACTAAAAAAGTGAACGCGCA 640

```

Db 10117 ABRSEMTARGAAGAAATSYSGNHAATSTASGTHRARGSRASGAGYGGASNGTHRHVAAGSGARG 10176
QY 641 GCGGCGGAAAGAAACAGGCTTTCAGACGAATGCGAAACACAGAGGCGCGGACGCGGACGC 700
Db 10177 HBRARGGASNSRSRGNVATHTHRHSRCYSGYHSASGNASNGYCTRYTASNGYSGHTH 10236
QY 701 GACCCCGGCGGAGAGGAGCCCACTGGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 760
Db 10237 RASYSASYSRARGSGNYSVAGARGYSGNRAATHTASARGSRMTASNGNGSRGY 10296
QY 761 GAGCCACACGCGCGCGCGCTGATGAGACGAGACGCGCGGAGCGGAGCGGAGCGGAGGAGG 820
Db 10297 STHRASRARGGNSRHYSRSGASNGYGNASNGYSASNGYSASNSRASNASNGNHTH 10356
QY 821 GCGGCGGAGAGGAGGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 879
Db 10357 RGAAGYSAAAHASNSYVASHSRGGASYSASNSRASNYSGGGAAAYSGYSGGNGSGR 10416
QY 880 CGGCGGAGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 939
Db 10417 MHSRYGCGYSRSHSASYSASYSRARGASNASTRGTHRAAASRGYSAAGCYSGNGT 10476
QY 940 CAGCAGAGCGCCACCGA 954
Db 10477 HRASNGYSGNYSAA 10491

RESULT 8
ADA71938
ID ADA71938 standard, DNA, 2000 BP.
XX
AC ADA71938;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5263.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001, 2001WO-1B001105.
XX
PR 22-JUN-2001, 2001WO-1B001105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX MPI, 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 27; SEQ ID NO 5263; 899bp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to

```

```

CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 2000 BP, 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;
Query Match 5.2%; Score 49.6; DB 7; Length 2000;
Best Local Similarity 8.7%; Pred. No. 0.0058;
Matches 40; Conservative 217; Mismatches 202; Indels 0; Gaps 0;
QY 491 ATGATTACACACCGGTGTGCGGCGGACANCTAGATGATATTGCGCCGGATACACA 550
Db 75 RTGSKSSGSSYSGKKMKKRYKRSKRWGRGRGMRSPRRWGRYRCARSGMAGSGSR 134
QY 551 CCATGATTCCCGCGTGTGATGATGCGGATACCGGATCAATTCGCGCGG 610
Db 135 MMGKSRMYWYMCYRAGGSCCKRKSCKGSGMGTCTRRARSGSGMSAKTKSGMSGR 194
QY 611 AGCTACAGACTTAAATAATTGACCGCGGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 670
Db 195 MMMSGCRSGCRSRASYSRYGTSRYGTYKQWYTSABRCRAYWTTTSYMACSYTW 254
QY 671 CCAAACAGAGGCGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 730
Db 255 RSKRSMMWMMKMKMSRYSYMSYSYMMCTAKYSYSRWCTYTRGGGGRGATRW 314
QY 731 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 790
Db 315 GREGYMRMAMMYKYWYVYGYGKMGKMGWAGMMWMSRWKAKCYWKMWRMRYTRRR 374
QY 791 AGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 850
Db 375 WAKKSRTSRKRYKMKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYK 434
QY 851 GCGCCTAGGCGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 910
Db 435 RMKSYGEMRKMSKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYK 494
QY 911 GACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 949
Db 495 GMRKCRRRRGMRYMRKRYMSARYMTARCARKYS 533

RESULT 9
AAK53491/C
ID AAK53491 standard, DNA, 114955 BP.
XX
AC AAK53491;
XX
DT 05-JUL-1999 (first entry)
XX
DE Human adenosine A1 receptor antisense oligonucleotide fragment.
XX
KW Antisense oligonucleotide; multiple target; antisense treatment;
KW impaired respiration; inflammation; lung disease;
KW pulmonary vasoconstriction; inflammation; allergic rhinitis;
KW acute asthma; allergy; asthma; impeded respiration;
KW respiratory distress syndrome; pain; cystic fibrosis;
KW pulmonary hypertension; pulmonary vasoconstriction; emphysema;
KW chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
KW colon cancer; breast cancer; lung cancer; pancreatic cancer;
KW hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
KW prostate cancer; ss.
XX
OS Synthetic.
XX
PN WO9913886-A1.
XX
PD 25-MAR-1999.
XX
PF 17-SEP-1998; 98WO-US019419.
XX
PR 17-SEP-1997; 97US-0059160P.
PR 09-JUN-1998; 98US-00093972.

```


PF	11-OCT-2001; 2001US-00981353.
XX	
PR	11-OCT-2000; 2000US-0239841P.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 15:48:07 ; Search time 79.981 Seconds
(without alignments)
6640.180 Million cell updates/sec

Title: US-10-076-747-7
Perfect score: 957
Sequence: 1 ttg9gcttactaatgcatg.....ggcagcagagacacgcagct 957

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 662709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	53.8	5.6	7218	1 US-08-232-463-14	Sequence 14, Appl
C 2	42.8	4.5	906	4 US-09-252-991A-1329	Sequence 1329, Ap
C 3	42.8	4.5	1023	4 US-09-252-991A-1283	Sequence 1283, Ap
C 4	42.8	4.5	1080	4 US-09-252-991A-1187	Sequence 1187, Ap
C 5	42.4	4.4	8091	4 US-09-230-652-1	Sequence 1, Appl
C 6	42.4	4.4	8257	4 US-09-484-970B-65	Sequence 65, Appl
C 7	41.8	4.4	3772	4 US-09-962-665-5	Sequence 5, Appl
C 8	41.6	4.4	3772	4 US-09-963-333-5	Sequence 3, Appl
C 9	41.6	4.3	4488	3 US-08-406-030A-3	Sequence 11532, A
C 10	41.2	4.3	417	4 US-09-252-991A-11532	Sequence 11531, A
C 11	41.2	4.3	1194	4 US-09-252-991A-11851	Sequence 11851, A
C 12	41.2	4.3	1485	4 US-09-252-991A-11591	Sequence 17, Appl
C 13	40.2	4.2	289	3 US-09-007-005-17	Sequence 17, Appl
C 14	40.2	4.2	289	3 US-09-244-796-17	Sequence 932, App
C 15	40.2	4.2	2376	4 US-09-023-655-932	Sequence 1, Appl
C 16	40.2	4.2	2404	4 US-08-945-771-1	Sequence 10534, A
C 17	39.8	4.2	678	4 US-09-252-991A-10534	Sequence 10529, A
C 18	39.8	4.2	1890	4 US-09-252-991A-10529	Sequence 1, Appl
C 19	39.6	4.1	4523	4 US-08-473-716-1	Sequence 1, Appl
C 20	39.4	4.1	1430	4 US-09-492-985-1	Sequence 11, Appl
C 21	39.2	4.1	3241	4 US-09-434-288-11	Sequence 10, Appl
C 22	39.2	4.1	6453	1 US-08-306-691B-14	Sequence 14, Appl
C 23	39.2	4.1	6453	3 US-09-209-668-10	Sequence 8, Appl
C 24	39.2	4.1	6453	3 US-09-356-952-8	Sequence 10866, A
C 25	38.8	4.0	6757	4 US-08-621-976-10866	Sequence 2, Appl
C 26	38.6	4.0	4403	765 3 US-09-103-840A-2	
C 27	38.6	4.0	4403	765 3 US-09-103-840A-2	

28	38.2	4.0	498	4 US-09-252-991A-3733	Sequence 3733, Ap
29	38.2	4.0	987	3 US-09-159-106-12	Sequence 12, Appl
30	38.2	4.0	1092	4 US-09-252-991A-3826	Sequence 3826, Ap
31	38.2	4.0	1516	3 US-09-159-106-10	Sequence 10, Appl
32	38.2	4.0	2112	4 US-09-252-991A-3782	Sequence 3782, Ap
33	38.2	4.0	3841	3 US-08-888-077A-33	Sequence 33, Appl
34	38.2	4.0	30001	1 US-08-125-468-1	Sequence 1, Appl
35	38.2	4.0	30001	2 US-08-474-933-1	Sequence 169, App
36	37.8	3.9	408	4 US-09-025-769B-169	Sequence 1031, Ap
37	37.8	3.9	489	4 US-09-252-991A-1031	Sequence 3, Appl
38	37.8	3.9	705	4 US-09-372-425A-3	Sequence 13250, A
39	37.8	3.9	840	4 US-09-252-991A-13250	Sequence 23, Appl
40	37.8	3.9	891	3 US-09-478-672A-23	Sequence 11, Appl
41	37.8	3.9	902	2 US-08-378-939-11	Sequence 12929, A
42	37.8	3.9	909	4 US-09-252-991A-12929	Sequence 1074, Ap
43	37.8	3.9	1083	4 US-09-252-991A-1074	Sequence 1077, Ap
44	37.8	3.9	1257	4 US-09-252-991A-1027	Sequence 978, App
45	37.8	3.9	1692	4 US-09-252-991A-978	

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHETFLINGER, F.
APPLICANT: PARKER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22113-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29, 766
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F1s
US-08-232-463-14
Query Match 5.6%, Score 53.8, DB 1, Length 7218;

US-09-252-991A-1187

Query Match 4.5%; Score 42.8; DB 4; Length 1080;
 Best Local Similarity 47.7%; Pred. No. 0.016;
 Matches 125; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 687 GCAGACCGGCGAGCCACCCCGGCGAGAGCCCACTGCGCGAGGCGGAGAGGAC 746
 DB 513 GCCGACCAACAGAGGAGTCAAGCCCGCGACATGCTGCTGATGCGAGAACACCCC 454
 QY 747 AGATACGAGGAGCGCGAGCCACACGCGCGCGTGTATAGACGAGAGCGCGAGGAGC 806
 DB 453 GCCCGAGTACAGGCGCATGATGACCAAGCGCGGCGACCGCGACATGCGCGAGCCGTGC 394
 QY 807 GCAGACCGGAGACCGCGCGCGCAAGACGACCGCGCAACACCGCGCTTGGCG 866
 DB 393 CGAGCCCGGCGCGAGCGCGACCGCTGAGCAGGCGGAGCGCGCGAGGCGCGGAG 334
 QY 867 CCAAGCGCGCGCGCGCGAGGCGCGCGCGAGACCGCGAGAGACCGACAGACCC 926
 DB 333 GAGCAGGCGGATGCGCGAGAGGTGAAGGCGCTGCGCGTACGCGAGCGCGAATTGCGCGC 274
 QY 927 GACCGCGCGAGCGCGAGAGC 948
 DB 273 GAGGAATCCAGTCTCCAGAGC 252

RESULT 5

US-09-230-652-1/c
 ; Sequence 1, Application US/09230652A
 ; Patent No. 6537775

GENERAL INFORMATION:
 APPLICANT: Tournier-Lasserre, Elisabeth
 APPLICANT: Jouteil, Anne
 APPLICANT: Boussez, Marie-Germaine
 APPLICANT: Bach, Jean-Francois
 TITLE OF INVENTION: GENE INVOLVED IN CADASIL, METHOD OF DIAGNOSIS AND
 FILE REFERENCE: 03715.0048-00000
 CURRENT APPLICATION NUMBER: US/09/230.652A
 EARLIER FILING DATE: 1999-05-17
 EARLIER APPLICATION NUMBER: FR 96 09733
 EARLIER FILING DATE: 1996-08-01
 EARLIER APPLICATION NUMBER: FR 97 04680
 EARLIER FILING DATE: 1997-04-16
 EARLIER APPLICATION NUMBER: PCT/FR97/01433
 EARLIER FILING DATE: 1997-07-31
 NUMBER OF SEQ ID NOS: 163
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1
 LENGTH: 8091
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (79)..(7041)
 OTHER INFORMATION: human ADYC No. 653775sch 3
 US-09-230-652-1

Query Match 4.4%; Score 42.4; DB 4; Length 8091;
 Best Local Similarity 47.7%; Pred. No. 0.045;
 Matches 124; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 675 ACACAGAGGCGCGAGACCGCGAGCGACCCCGGCGAGAGGAGCCCACTGCGCGAGGCG 734
 DB 4112 AGACAGGCGGCGCGCGCGCACTGCTGTGCGCCCGCGCGAGACCCCGGAGAGCTG 4053
 QY 735 GAGCGGAAGACAGATACAGAGACCGCGACCGCGCGCGCTGTGATGAGACGAGAC 794
 DB 4052 CGGCGAGAGGTCGCGACCACTGCGGCGGAGCGAGCGAGCGCGCGCGGCGTCTGC 3993
 QY 795 GCGCGAGGAGCGAGACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 854

DB 3992 TGCGATGAGGACGCGCCACCGGAGACTGCACTCCCGGACAGAGAGCGCGCACCCGCTCGCA 3933
 QY 855 CTAGGCGCTGCGCGACAGCGCGCGCGAGAGCGCGGCGGAGAGCACCGGAGAGAGC 914
 DB 3932 CGCGGACCCCGAGACCGCTGCGGACAGTACAGGTAAAGTACGCCACCCCGAGAGCC 3873
 QY 915 GACCAAGGACCGGACCGGCG 934
 DB 3872 GGGCTAGAGCGGCACTGGCC 3853

RESULT 6

US-09-484-970B-65/c
 ; Sequence 65, Application US/09484970B
 ; Patent No. 6426186

GENERAL INFORMATION:
 APPLICANT: Jones, Karen A.
 APPLICANT: Volkuth, Wayne
 APPLICANT: Walker, Michael G.
 TITLE OF INVENTION: BONE REMODELING GENES
 FILE REFERENCE: PB-0014 US
 CURRENT APPLICATION NUMBER: US/09/484,970B
 NUMBER OF SEQ ID NOS: 172
 SOFTWARE: PERL Program
 SEQ ID NO 65
 LENGTH: 8257
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 OTHER INFORMATION: Incyte ID No. 6426186 285840.2
 NAME/KEY: unsure
 LOCATION: 7391, 8247
 OTHER INFORMATION: a, t, c, g, or other
 US-09-484-970B-65

Query Match 4.4%; Score 42.4; DB 4; Length 8257;
 Best Local Similarity 47.7%; Pred. No. 0.045;
 Matches 124; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 675 ACACAGAGGCGCGAGACCGCGAGCGACCCCGGCGAGAGGAGCCCACTGCGCGAGGCG 734
 DB 4112 AGACAGGCGGCGCGCGCGCACTGCTGTGCGCCCGCGGAGAGCACCGGAGAGCTG 4053
 QY 735 GAGCGGAAGACAGATACAGAGACCGCGACCGCGCGCGCTGTGATGAGACGAGAC 794
 DB 4052 CGGCGAGAGGTCGCGGACCACTGAGGCGGCGAGCGCGCGCGCGCGCGCTTGC 3993
 QY 795 GCGCGAGGAGCGCGAGACCCCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 854
 DB 3992 TGCGATGAGGACCGCGACCGCGGCACTGCACTCCCGCGAGAGCGCGCGCGCGCG 914
 QY 855 CTAGGCGCTGCGCGACAGCGCGCGCGCGAGAGCGCGGCGGAGAGCACCGGAGAGAGC 914
 DB 3932 CGGCGACCCCGAGAGCGCTGCGGACAGTACAGGTAAAGTACGCCACCCCGAGAGCC 3873
 QY 915 GACCAAGGACCGGACCGGCG 934
 DB 3872 GGGCTAGAGCGGCACTGGCC 3853

RESULT 7

US-09-962-665-5
 ; Sequence 5, Application US/09962665
 ; Patent No. 6537759

GENERAL INFORMATION:
 APPLICANT: Stanton, Jr., Vincent P.
 TITLE OF INVENTION: POLYPOLYGLUTAMATE SYNTHETASE GENE SEQUENCE
 TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
 FILE REFERENCE: 11926-015004
 CURRENT APPLICATION NUMBER: US/09/962.665

;; CURRENT FILING DATE: 2001-09-24
;; PRIOR APPLICATION NUMBER: 09/658,659
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 09/596,033
;; PRIOR FILING DATE: 2000-06-15
;; PRIOR APPLICATION NUMBER: 09/357,743
;; PRIOR FILING DATE: 1999-07-20
;; PRIOR APPLICATION NUMBER: 09/357,024
;; PRIOR FILING DATE: 1999-07-19
;; PRIOR APPLICATION NUMBER: 60/093,484
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 5
;; LENGTH: 3772
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 431, 441
;; OTHER INFORMATION: n = a or g
;; NAME/KEY: misc_feature
;; LOCATION: 498
;; OTHER INFORMATION: n = c or t
;; NAME/KEY: misc_feature
;; LOCATION: 579, 599
;; OTHER INFORMATION: n = g or c
US-09-963-333-5

Query Match 4.4%; Score 41.8; DB 4; Length 3772;
Best Local Similarity 50.8%; Pred. No. 0.049;
Matches 153; Conservative 0; Mismatches 142; Indels 6; Gaps 2;

QY 660 TCAGCGGATGCGCAACACAGAGGCGCGCGAGCCCGCGCGAGAGAGCC 719
DB 1041 TCAGGGGGGGGGCGCGCGAGGCGTGAAGCGAGAGGCGCGTGAAC 1100
QY 720 CCACTGCGCGCAGGCGCGAGCGCAAGATACAGAGACCGAGCCCGCGCGT 779
DB 1101 CGGCTGGGAGAGAGCGCGCGCGCGCGCGTGAAGAGCGCGTGGAGGTC 1160
QY 780 GCATGAGACGAGAGAGCGCGCGAGAGCGCGAGCGCGCGCGCGCGCGCG 839
DB 1161 G-GGAGCGTCTCTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCTGTA 1219
QY 840 GCAAGCCACACGCGCGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 899
DB 1220 GCGGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTA 1279
QY 900 GCAAGCGAG-----GAGACCGACACAGCGCGCGCGCGCGCGCGCGCG 954
DB 1280 GTCCCGAGATCCGCGTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1339
QY 955 G 955
DB 1340 G 1340

RESULT 8
US-09-963-333-5
;; Sequence 5, Application US/0996333
;; Patent No. 6664062
;; GENERAL INFORMATION:
;; APPLICANT: Stanton, Jr., Vincent P.
;; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
;; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
;; FILE REFERENCE: 11926-015002
;; CURRENT APPLICATION NUMBER: US/09/963,333
;; CURRENT FILING DATE: 2001-09-24
;; PRIOR APPLICATION NUMBER: 09/658,659
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 09/596,033

;; PRIOR FILING DATE: 2000-06-15
;; PRIOR APPLICATION NUMBER: 09/357,743
;; PRIOR FILING DATE: 1999-07-20
;; PRIOR APPLICATION NUMBER: 09/357,024
;; PRIOR FILING DATE: 1999-07-19
;; PRIOR APPLICATION NUMBER: 60/093,484
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 5
;; LENGTH: 3772
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 431, 441
;; OTHER INFORMATION: n = a or g
;; NAME/KEY: misc_feature
;; LOCATION: 498
;; OTHER INFORMATION: n = c or t
;; NAME/KEY: misc_feature
;; LOCATION: 579, 599
;; OTHER INFORMATION: n = g or c
US-09-963-333-5

Query Match 4.4%; Score 41.8; DB 4; Length 3772;
Best Local Similarity 50.8%; Pred. No. 0.049;
Matches 153; Conservative 0; Mismatches 142; Indels 6; Gaps 2;

QY 660 TCAGCGGATGCGCAACACAGAGGCGCGCGAGCCCGCGCGAGAGAGCC 719
DB 1041 TCAGGGGGGGGGCGCGCGAGGCGTGAAGCGAGAGGCGCGTGAAC 1100
QY 720 CCACTGCGCGCAGGCGCGAGCGCAAGATACAGAGACCGAGCCCGCGCGT 779
DB 1101 CGGCTGGGAGAGAGCGCGCGCGCGCGCGTGAAGAGCGCGTGGAGGTC 1160
QY 780 GCATGAGACGAGAGAGCGCGCGAGAGCGCGAGCGCGCGCGCGCGCGCG 839
DB 1161 G-GGAGCGTCTCTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCTGTA 1219
QY 840 GCAAGCCACACGCGCGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 899
DB 1220 GCGGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTA 1279
QY 900 GCAAGCGAG-----GAGACCGACACAGCGCGCGCGCGCGCGCGCGCG 954
DB 1280 GTCCCGAGATCCGCGTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1339
QY 955 G 955
DB 1340 G 1340

RESULT 9
US-08-406-030A-3
;; Sequence 3, Application US/08406030A
;; Patent No. 6270985
;; GENERAL INFORMATION:
;; APPLICANT: Treco, Douglas A.
;; APPLICANT: Heartlein, Michael W.
;; APPLICANT: Haug, Brian M.
;; APPLICANT: Seiden, Richard P.
;; TITLE OF INVENTION: Protein Production and Delivery
;; NUMBER OF SEQUENCES: 30
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
;; STREET: Two Militia Drive
;; CITY: Lexington
;; STATE: Massachusetts
;; COUNTRY: USA
;; ZIP: 02173
;; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/406,030A
FILING DATE: 17-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/243,391
FILING DATE: 13-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,586
FILING DATE: 03-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/911,533
FILING DATE: 10-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,840
FILING DATE: 05-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/789,188
FILING DATE: 05-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/11704
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/09627
FILING DATE: 05-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: TKT95-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4488 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-406-030A-3

Query Match 4.3%; Score 41.6; DB 3; Length 4488;
Best Local Similarity 48.0%; Pred. No. 0.06;
Matches 190; Conservative 1; Mismatches 191; Indels 14; Gaps 2;

QY 563 CCCCGTGGCTGATGGAATATGCGGTACGCGGCATCAATTGCGCCGGAGGCTACAGACCT 622
DB 2206 CTCCTTCTCAGAACGATGCGGACGACGCGGCTTCGGGCGCGCGGGAGGCCCA 2265
QY 623 AAAAAATTGACCGCGCAGCGCGCGGGAAGAACAGGCTTTCGACGGAATGCCAAACACAGAG 682
DB 2266 GCACACACAGCCCGAGAGCAGCAGCGCGGCGCGCGGAGGCGCH-----C 2316
QY 683 GCGCGCAGACCGCGAGCGGCGACCCCGCGCGGAGAGCCCACTCGCGCAGCGCGAGCGCAA 742
DB 2317 GGCATGACCGGAACGGGACAGCTGGGAGGAGGAGGAGGAGGAGGCGCGGAGCGGGCGG 2376
QY 743 GACAGATACGAGAGCAGCGACCAACGCGCGCGCTGATGAGCAGGAGCGCGCGAGG 802
DB 2377 AGGAGAGGAGCGCGGAGTGGCGAGGCGAGGCGCGCGCGCGGCGGCGGCGGAGG 2436
QY 803 GAGCGCAGACCCGAGAGCGCGCGCAAGAGCGACCGG-----GCAAGCCACAGCGCCCTTA 857
DB 2437 GGGCGGGGCGCGGGCGCGGGCGCGGCGCACTGCGCGGAGGCGCTGCGGCGCGCGCA 2496
QY 858 GGCCTTCGCGCAGCGCGCGCGCAGCGCGGAGGCGGGGCGGAGCAGCGACCGAGCAGCAG 917
DB 2497 GAGTCGCGCGCGCGCGCGCGCGCGGCGGAGGAGCGCGCGGAGGAGCGCGCGCGCGGCGG 2556

QY 918 CACGAGCCCGACCGCGCGCGGAGGAGCAGACGCCACCG 953
DB 2557 CGCTGACCGCGCGCGGTACGCGGCTCTACTGCGCCCG 2592

RESULT 10
US-09-252-991A-11532/c
Sequence 11532, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 11532
LENGTH: 417
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11532

Query Match 4.3%; Score 41.2; DB 4; Length 417;
Best Local Similarity 50.0%; Pred. No. 0.031;
Matches 103; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 710 CCGAGAGACCCCACTGCGCGCGGAGGCGGAGGAGCAGATACGAGACCGGAGCCACAC 769
DB 359 CGTCATGCTCTGCGCAGGAGCGCGGCGCTTCGCTGCGCAGCCATACGACCGGACCTTGC 300
QY 770 GCGCGCCGCTGATGAGCAGGAGCGCGCGGAGGAGCAGACCGCAAGCAGCGCGCCAA 829
DB 299 GCAATCCGCGAGTCGCGCGCTACGCGCATGCGCATGCGCATGCGCATGCGCATGCG 240
QY 830 GAGCGACCGCGCAAGCCACGACGCGCTTAGGCGCTTGCAGCGCGCGCGCGCGAGAG 889
DB 239 GCGCGCCAGCG 180
QY 890 GCGCGCGCGAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 915
DB 179 CCGCGCGGCTTCCACTGCGCGCAACG 154

RESULT 11
US-09-252-991A-11851
Sequence 11851, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 11851
LENGTH: 1194
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11851

Query Match 4.3%; Score 41.2; DB 4; Length 1194;
Best Local Similarity 50.0%; Pred. No. 0.046;
Matches 103; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

```

QY 710 CGGAGAGCCCACTGCGGAGGCGGAGCCGAGACGATACGAGACCGGACCCAC 769
    |||
Db 687 CCGTCAATGTCGCGGAGACCGGCGCTCTGGTGCCACCATACGACGCGACCTCGC 746
QY 770 GCGCGCCCGTCATGAGACGAGACGCGCGGAGGAGCGGACCGGACCGGCGCAA 829
    |||
Db 747 GCATGTCGCGGAGTCGCGGCGCTACGCGATGAGCCATCGGAGTCCGACCACTGGA 806
QY 830 GAGGACCGCGGAGACCGGACCGGCTAGGCGCTCGGCGGACGCGGCGGCGGAG 889
    |||
Db 807 GCGCCGCGGAGCGACCGACGAGACGCGGCTGAGTGTGATGAGCGCGGCGCAACGTGT 866
QY 890 GCGGCGCGGAGACCGGACCGGAGACCG 915
    |||
Db 867 CCGCGCGGCTTCCCACTCGGCGGACG 892

```

```

RESULT 12
US-09-252-991A-11591/C
; Sequence 11591, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11591
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11591

```

```

Query Match      4.3%; Score 41.2; DB 4; Length 1485;
Best Local Similarity 50.0%; Pred. No. 0.051;
Matches 103; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

```

```

QY 710 CGGAGAGCCCACTGCGGAGGCGGAGCCGAGACGATACGAGACCGGACCCAC 769
    |||
Db 622 CCGTCAATGTCGCGGAGACCGGCGCTCTGGTGCCACCATACGACGCGACCTCGC 563
QY 770 GCGCGCCCGTCATGAGACGAGACGCGCGGAGGAGCGGACCGGACCGGCGCAA 829
    |||
Db 562 GCATGTCGCGGAGTCGCGGCGCTACGCGATGAGCCATCGCGAGTCCGACCACTGGA 503
QY 830 GAGGACCGCGGAGACCGGACCGGCTAGGCGCTCGGCGGACGCGGCGGCGGAG 889
    |||
Db 502 GCGCCGCGGAGCGGACCGGACCGGCTAGGCGCTGAGTGTGATGAGCGGCGGCGCAA 443
QY 890 GCGGCGCGGAGACCGGACCGGAGACCG 915
    |||
Db 442 CCGCGCGGCTTCCCACTCGGCGGACG 417

```

```

RESULT 13
US-09-007-005-17
; Sequence 17, Application US/0907005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihne
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B

```

```

; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; NAME/KEY: misc_feature
; LOCATION: (1)-(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

```

```

Query Match      4.2%; Score 40.2; DB 3; Length 289;
Best Local Similarity 6.7%; Pred. No. 0.051;
Matches 17; Conservative 109; Mismatches 127; Indels 0; Gaps 0;

```

```

QY 608 GCGAGCTACAGACTTAAAAAGTTGACCGCGGAGCGGCGGCGGAGACGAGCTTCAGCGA 667
    |||
Db 2 GCGGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 61
QY 668 ATGCCAACACAGAGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 727
    |||
Db 62 SRRNRSSRRNRSSRRNRSSRRNRSSRRNRSSRRNRSSRRNRSSRRNRSSRRNRSSRRNR 121
QY 728 GCGAGCGGAGCGGAGGACGATACGAGACCGGACCGGACCGGCGGCGGCGGAGTGA 787
    |||
Db 122 SRRNRSSRRNRSSRRNRSSRRNRSSRRNRSSRRNRSSRRNRSSRRNRSSRRNRSSRRNR 181
QY 788 CGGAGCGGCGGAGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 847
    |||
Db 182 SRRNRSSRRNRSSRRNRSSRRNRSSRRNRSSRRNRSSRRNRSSRRNRSSRRNRSSRRNR 241
QY 848 CACGCGCTAGGC 860
    |||
Db 242 CRURCRURGR 254

```

```

RESULT 14
US-09-244-796-17
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihne
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350007
; CURRENT APPLICATION NUMBER: US/09/244,796
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 09/007,005
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; NAME/KEY: misc_feature

```

LOCATION: (1)...(289)
OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

Query Match 4.2%; Score 40.2; DB 3; Length 289;
Best Local Similarity 6.7%; Pred. No. 0.051;
Matches 17; Conservative 109; Mismatches 127; Indels 0; Gaps 0;

QY GGGAGCTACAGACCTAAAGTTGACCGCGCGCGGCGGAGAGCGGCTTCGACCGA 667
DB 2 GGGGAC 61
QY 668 ATGCCAAACACAGAGCGCGGAGACCGCGCGCGGCGGAGAGCGGCTTCG 727
DB 62 SRNRSSNRSSNRSSNRSSNRSSNRSSNRSSNRSSNRSSNRSSNRSSNRSSNR 121
QY 728 GCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 787
DB 122 SRNRSSNRSSNRSSNRSSNRSSNRSSNRSSNRSSNRSSNRSSNRSSNRSSNR 181
QY 788 CGAGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 847
DB 182 SRNRSSNRSSNRSSNRSSNRSSNRSSNRSSNRSSNRSSNRSSNRSSNRSSNR 241
QY 848 CACGCGCTAGGC 860
DB 242 CRURCRURGRC 254

RESULT 15
US-09-023-655-932
Sequence 932, Application US/09023655
Patent No. 6607879

GENERAL INFORMATION:

APPLICANT: Cooke, Benjamin G.

APPLICANT: Jeffrey J. Sellman

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

TITLE OF INVENTION: EXPRESSION

NUMBER OF SEQUENCES: 1508

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/023,655

FILING DATE: HEREMITH

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

TELECOMMUNICATION INFORMATION:

REFERENCE/DOCKET NUMBER: PA-0001 US

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 932:

SEQUENCE CHARACTERISTICS:

LENGTH: 2376 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: 91369836
US-09-023-655-932

Query Match 4.2%; Score 40; DB 4; Length 2376;
Best Local Similarity 47.0%; Pred. No. 0.13;
Matches 124; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 689 AGACCGGACGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 748
DB 54 AGCCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 113
QY 749 ATACGAGGACGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 808
DB 114 ACCCTGGGCTCAGCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 173
QY 809 AGACCGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 868
DB 174 GTCCCGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 233
QY 869 AGCGCGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 928
DB 234 GGGGCGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 293
QY 929 CCGGCGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 952
DB 294 CGGCGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 317

Search completed: April 26, 2004, 20:28:38
Job time: 83.961 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 18:46:58 : Search time 398.717 Seconds
(without alignments)
10821.600 Million cell updates/sec

Title: US-10-076-747-7
Perfect score: 957
Sequence: 1 ttgggcttactaatgcatgct.....ggcagcagacccacgagct 957

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2907579 seqs, 2254313464 residues
Total number of hits satisfying chosen parameters: 5815158

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PC1_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PC1US_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10C_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	956	99.9	957	15	US-10-076-747-7 Sequence 7, Appl1
2	809.6	84.6	1460	15	US-10-076-747-8 Sequence 8, Appl1
3	55.2	5.8	349	15	US-10-074-475-147 Sequence 147, Appl
4	52.4	5.5	594	13	US-10-142-426-10 Sequence 10, Appl
5	52.4	5.5	594	15	US-10-123-155-10 Sequence 10, Appl
6	52.4	5.5	594	15	US-10-146-731-10 Sequence 10, Appl
7	52.4	5.5	594	15	US-10-140-472-10 Sequence 10, Appl
8	52.4	5.5	594	15	US-10-141-761-10 Sequence 10, Appl
9	52.4	5.5	594	15	US-10-142-885-10 Sequence 10, Appl
10	52.4	5.5	594	15	US-10-158-790-10 Sequence 10, Appl
11	52.4	5.5	594	15	US-10-137-871-10 Sequence 10, Appl
12	52.4	5.5	594	16	US-10-140-823-10 Sequence 10, Appl
13	52.4	5.5	594	16	US-10-141-756-10 Sequence 10, Appl
14	52.4	5.5	594	16	US-10-141-759-10 Sequence 10, Appl

15	52.4	5.5	594	16	US-10-140-805-10	Sequence 10, Appl
16	52.4	5.5	594	16	US-10-140-864-10	Sequence 10, Appl
17	52.4	5.5	365	13	US-10-425-114-24362	Sequence 34362, A
18	50.2	5.2	872	15	US-10-198-846-6247	Sequence 6247, Ap
19	44.4	4.6	1185	13	US-10-425-114-2918	Sequence 2918, Ap
20	44.4	4.6	2265	16	US-10-062-674-1761	Sequence 1761, Ap
21	43.2	4.5	3600	15	US-10-198-846-13029	Sequence 13029, A
22	43.2	4.5	71292	13	US-10-087-192-1942	Sequence 1942, Ap
23	43	4.5	366	9	US-09-960-353-9109	Sequence 9109, Ap
24	43	4.5	764	9	US-09-981-353-46	Sequence 46, Appl
25	42.8	4.5	5760	13	US-10-152-886-14	Sequence 14, Appl
26	42.4	4.4	2754	15	US-10-156-761-2167	Sequence 2167, Ap
27	42.4	4.4	2919	13	US-10-282-122A-14574	Sequence 14574, A
28	42.4	4.4	8091	15	US-10-101-510-86	Sequence 86, Appl
29	42.4	4.4	8091	15	US-10-356-665-1	Sequence 1, Appl1
30	42.4	4.4	8091	16	US-10-159-563-322	Sequence 322, App
31	42.4	4.4	9025608	15	US-10-156-761-1	Sequence 1, Appl1
32	42.2	4.4	895	15	US-10-078-090-78	Sequence 78, Appl
33	41.2	4.3	1440	15	US-10-156-761-6190	Sequence 6190, Ap
34	41.2	4.3	1164	9	US-09-815-242-4078	Sequence 4078, Ap
35	41.2	4.3	1164	13	US-10-282-122A-7370	Sequence 7370, Ap
36	41.2	4.3	3066	15	US-10-121-988-152	Sequence 152, App
37	41.2	4.3	3066	15	US-10-200-562-152	Sequence 152, App
38	41.2	4.3	3066	15	US-10-237-551-152	Sequence 152, App
39	41.2	4.3	154746	10	US-09-827-688-8	Sequence 8, Appl1
40	41.2	4.3	154746	10	US-09-827-688-8	Sequence 8, Appl1
41	41	4.3	984	13	US-10-282-122A-25620	Sequence 25620, A
42	40.6	4.2	536	17	US-10-338-110-119	Sequence 119, App
43	40.6	4.2	544	13	US-10-424-599-973	Sequence 973, App
44	40.6	4.2	721	13	US-10-425-114-18851	Sequence 18851, A
45	40.6	4.2	2277	15	US-10-156-761-3085	Sequence 3085, Ap

ALIGNMENTS

RESULT 1
US-10-076-747-7 Application US/10076747
Publication No. US20030180726A1
GENERAL INFORMATION:
APPLICANT: Salceda, Susana
APPLICANT: Macina, Roberto
APPLICANT: Hu, Ping
APPLICANT: Recipon, Heve
APPLICANT: Kaira, Kalpana
APPLICANT: Caffarely, Robert
APPLICANT: Sun, Yongming
APPLICANT: Liu, Chenghua
TITLE OF INVENTION: Compositions and Methods Relating to Ovarian Specific Genes and
FILE REFERENCE: DEX-0315
CURRENT APPLICATION NUMBER: US/10/076,747
CURRENT FILING DATE: 2002-02-13
PRIOR APPLICATION NUMBER: 60/268,290
PRIOR FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: 60/268,834
NUMBER OF SEQ ID NOS: 129
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 957
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (519)..(519)
OTHER INFORMATION: a, c, g or t
US-10-076-747-7

Query Match 99.9% Score 956, DB 15; Length 957;
Best Local Similarity 100.0%; Pred. No. 1.8e-293;
Matches 957; Conservative 0; Mismatches 0; Gaps 0;

QY 1 TTGGGCTTACTAATGATGCTGAGCGGCGCCAGTGTGATGATGCTGTGTGCGCGC 60
 DB 1 TTGGGCTTACTAATGATGCTGAGCGGCGCCAGTGTGATGATGCTGTGTGCGCGC 60
 QY 61 CGAGGTACCGGGT CAGCAGCTATCTGAGCTTGA CGGCTGATG CAGTGTGAGTCCAC 120
 DB 61 CGAGGTACCGGGT CAGCAGCTATCTGAGCTTGA CGGCTGATG CAGTGTGAGTCCAC 120
 QY 121 AGAAGCTACAGCCTGACAGAGCTTATGCTGATGAGGAGACACCGGTGATGAGACA 180
 DB 121 AGAAGCTACAGCCTGACAGAGCTTATGCTGATGAGGAGACACCGGTGATGAGACA 180
 QY 181 GTGTGCTCCCTACAGAAATGTCTATAGGTTCTTAACAGCTCAGCCCCCACTACCAATG 240
 DB 181 GTGTGCTCCCTACAGAAATGTCTATAGGTTCTTAACAGCTCAGCCCCCACTACCAATG 240
 QY 241 GGGAGCTAGCAGCTGCGAGGATGCCAGAGGAGGGGTCTCTCCATCCACACCAACA 300
 DB 241 GGGAGCTAGCAGCTGCGAGGATGCCAGAGGAGGGGTCTCTCCATCCACACCAACA 300
 QY 301 AGGGCAGTCAAAAGCCTTATCATGCGCATGTCAGCGTATGTAAGAGCCTTACAAAT 360
 DB 301 AGGGCAGTCAAAAGCCTTATCATGCGCATGTCAGCGTATGTAAGAGCCTTACAAAT 360
 QY 361 AAGATTTCTGCACTTGGTTGAATGTCTTACATACATPACAAAGACATATCTCACTA 420
 DB 361 AAGATTTCTGCACTTGGTTGAATGTCTTACATACATPACAAAGACATATCTCACTA 420
 QY 421 CAGGAGTGCATGCATCACCAGTCCGTGGGCGAGATGCACTTGGCTCTGTGGGTCCA 480
 DB 421 CAGGAGTGCATGCATCACCAGTCCGTGGGCGAGATGCACTTGGCTCTGTGGGTCCA 480
 QY 481 ATGATGACTCATATGTTTACACACCGGTGTGGGCGACANCTATGATGATATTTGCCCG 540
 DB 481 ATGATGACTCATATGTTTACACACCGGTGTGGGCGACANCTATGATGATATTTGCCCG 540
 QY 541 GGAATCAGACCAATGATTTCCCGCGTGTGATGCAATATGCGGTACGCGCATCAA 600
 DB 541 GGAATCAGACCAATGATTTCCCGCGTGTGATGCAATATGCGGTACGCGCATCAA 600
 QY 601 TTCCGCGGAGCTACAGACTTAAAGTTGACCGCGAGCGGCGGAGAAACAGGCTTT 660
 DB 601 TTCCGCGGAGCTACAGACTTAAAGTTGACCGCGAGCGGCGGAGAAACAGGCTTT 660
 QY 661 CGAGCAATGCTCAACACAGAGGCGCGCAGACCGGCAAGCCGCCGCGAGAGGCC 720
 DB 661 CGAGCAATGCTCAACACAGAGGCGCGCAGACCGGCAAGCCGCCGCGAGAGGCC 720
 QY 721 CACTGCGGAGGCGGAGGAGACAGATACAGAGACGCGAGCCACCGCGCGCGCTG 780
 DB 721 CACTGCGGAGGCGGAGGAGACAGATACAGAGACGCGAGCCACCGCGCGCGCTG 780
 QY 781 CATGACCGAGACAGCGGCGGAGGAGCGAGACCGGCAAGCGCGCAAGAGCGAG 840
 DB 781 CATGACCGAGACAGCGGCGGAGGAGCGAGACCGGCAAGCGCGCAAGAGCGAG 840
 QY 841 CAAAGCACAAGGCGCTTATGAGCTTGGGCGACGCGCGGCGACGCGCGAGAGGCGGAG 900
 DB 841 CAAAGCACAAGGCGCTTATGAGCTTGGGCGACGCGCGGCGACGCGCGAGAGGCGGAG 900
 QY 901 CACCGCAGAGACCGACACAGAGACCGGACCGGCGGAGGAGCAGAGCCAGACT 957
 DB 901 CACCGCAGAGACCGACACAGAGACCGGACCGGCGGAGGAGCAGAGCCAGACT 957

RESULT 2

US-10-076-747-8

; Sequence 8, Application US/10076747

; Publication No. US20030180726A1

; GENERAL INFORMATION:

; APPLICANT: Salceda, Susana

; APPLICANT: Macina, Roberto

; APPLICANT: Hu, Ping

; APPLICANT: Recipon, Herve
 ; APPLICANT: Karia, Kalpana
 ; APPLICANT: Cafierkey, Robert
 ; APPLICANT: Sun, Yongming
 ; APPLICANT: Liu, Chenghua
 ; TITLE OF INVENTION: Compositions and Methods Relating to Ovarian Specific Genes and
 ; FILE REFERENCE: DEX-0315
 ; CURRENT APPLICATION NUMBER: US/10/076,747
 ; CURRENT FILING DATE: 2002-02-13
 ; PRIOR APPLICATION NUMBER: 60/268,290
 ; PRIOR FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/268,834
 ; NUMBER OF SEQ ID NOS: 129
 ; SOFTWARE: Patent version 3.1
 ; SEQ ID NO 8
 ; LENGTH: 1460
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1022)..(1022)
 ; OTHER INFORMATION: a, c, g or t
 US-10-076-747-8

Query Match: 84.6%; Score 809.6; DB 15; Length 1460;
 Best Local Similarity 96.5%; Pred. No. 8e-247;
 Matches 864; Conservative 0; Mismatches 19; Indels 12; Gaps 3;

QY 63 AGGTACGGGCTCAGCAGCTATCTGAGCTGAGCGGCTGATGAGTGTGAGTCCACAG 122
 DB 578 AAGTACGGGCTCAGCAGCTATCTGAGCTGAGCGGCTGATGAGTGTGAGTCCACAG 633
 QY 123 AAGCTACAGCTTGCACAGAGCTTATGCTGATGATGAGGAGACACCGTGTGAGACAGT 182
 DB 634 AAGCTACAGCTTGCACAGAGCTTATGCTGATGATGAGGAGACACCGTGTGAGACAGT 685
 QY 183 GGTTCCTTACAGAAATGTTCTATAGTCTCTTAAACGCTGAGCCCCCACTACCAATGGC 242
 DB 686 GGTTCCTTACAGAAATGTTCTATAGTCTCTTAAACGCTGAGCCCCCACTACCAATGGC 745
 QY 243 GAGACTAGCAGCTGCGAGGATCCCAAGGAGAGGGGTCTCTCATCCACACCAACAAG 302
 DB 746 GAGACTAGCAGCTGCGAGGATCCCAAGGAGAGGGGTCTCTCATCCACACCAACAAG 805
 QY 303 GGGAGTCAAAAGCCTTATCATGCGCATGTGACGCTATGTAAGAGCGCTTACAAATA 362
 DB 806 GGGAGTCAAAAGCCTTATCATGCGCATGTGACGCTATGTAAGAGCGCTTACAAATA 865
 QY 363 GATATCTGACCTTGTGTAATGTCTATCATATCAATTAACAAGACATATCACTACA 422
 DB 866 GATATCTGACCTTGTGTAATGTCTATCATATCAATTAACAAGACATATCACTACA 925
 QY 423 CGGAGTGCATGATCAACCGGTCCGTGCGGCGAATGCCATTGCTGTGTGCTCAAT 482
 DB 926 CGGAGTGCATGATCAACCGGTCCGTGCGGCGAATGCCATTGCTGTGTGCTCAAT 985
 QY 483 GATGACTCATATGTTACACACCGGTGTGCGGCGACANCTATGATGATATTCGCGGG 542
 DB 986 GATGACTCATATGTTACACACCGGTGTGCGGCGACANCTATGATGATATTCGCGGG 1045
 QY 543 ATCAGACAGCATGATTTCCCGCGTGTGATGCAATATGCGGTACGCGGATCAATT 602
 DB 1046 ATCAGACAGCATGATTTCCCGCGTGTGATGCAATATGCGGTACGCGGATCAATT 1105
 QY 603 CGCCCGGAGCTTACAGCTTAAAGTTGACCGGCGAGCGGCGGAGAAACAGGCTTTG 662
 DB 1106 CGCCCGGAGCTTACAGCTTAAAGTTGACCGGCGAGCGGCGGAGAAACAGGCTTTG 1165
 QY 663 ACGGAATGCCAAACACAGAGGCGCGGACGCGGACGCGGCGGAGAGAGGCCCA 722
 DB 1166 ACGGAATGCCAAACACAGAGGCGCGGACGCGGACGCGGCGGAGAGAGGCCCA 1225

macrines 27; conservative 105; misbehavior 205; models 0; gaps 0;

; Sequence 10, Application US/10146731

Query Match	5.5%;	Score 52.4;	DB 15,	Length 594;
Best Local Similarity	7.0%;	Pred. No. 5.2e-06;		
Matches	34;	Conservative 189;	Mismatches 265;	Indels 0;
			Gaps	0;

RESULT 7

US-10-140-472-10
; Sequence 10, Application US/10140472
; Publication No. US20030138888A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

```

/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3330R1C168
/ CURRENT APPLICATION NUMBER: US/10/140,472
/ PRIOR FILING DATE: 2002-05-06
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 10
/ LENGTH: 594
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-140-472-10

```

Query Match 5.5%; Score 52.4; DB 15; Length 594;
Best Local Similarity 7.0%; Pred. No. 5,2e-06;

Matches 34; Conservative 189; Mismatches 265; Indels 0; Gaps 0;

```

QY 393 ATACATAACAAGACACATCTCACTACAGAGAGATGATGATCACCAGTCCGTCGGCGG 452
DB 60 AEPVNMWPMQGDITNMIDRFVRAHLDHIDPYPLLTITTSPEQESDERKCNVERXRL 119
QY 453 CGAATGCCACTTCGCTCGTCGTCGTCATGATGATCACTCAACACACCGGTGCGG 512
DB 120 VQNDPAGISEBQCLYQIYIDELYGQLQRPSEDEKXLAEKASIGTYEDSVAVEKAA 179
QY 513 CGCACANTATGATGATGATTTGCGCGGATCAACAGCATGATTTCCCGCGTGGCT 572
DB 180 EKPEEESAAEESNSDEVDIVPIDVVDVDELNQVADLNKQATTYGMADGDFVRL 239
QY 573 GATGAAATATGCGGTACCGCGCATCAATTCGCCGGAGCTACAGACTTAAAGTTG 632
DB 240 RKDKEEABAIKHALEBEKAMYSGRSRRRQREFREKRLRGKTISSPYARSDPTYP 299
QY 633 ACCCGCGAGCGCGGAGAGAGAGCTTTCGACGGAATGCCAAACAGAGAGCGCGCAGAC 692
DB 300 YKRSPESSSSSRSSRSPTPGREGKITFTTSFGSDSEAAAAAAGVTTGKPPAP 359
QY 693 CGCGAGCGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 752
DB 360 PQGGPAPGRNARSARRSSSSSSSSASRTSSSSSSSSSSSSSSSSSSSSSSSSSSSS 419
QY 753 GAGACAGCGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 812
DB 420 RSMWSRSTRSRYSRSGRGRHSGGSRDGHYSRSPARRGGYPRRRSRSSHSQDRY 479
QY 813 CCGAAGCAGCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 872
DB 480 RRGGRGLHHSSSRSSWSLSPSRSLTRSRSHSPSPQSRSRSSRSGSPSPARE 539
QY 873 CCGGCCAC 880
DB 540 KLTRPAA 547

```

RESULT 8
US-10-141-761-10
Sequence 10, Application US/10141761

```

/ Publication No. US20030148432A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3330R1C198
/ CURRENT APPLICATION NUMBER: US/10/141,761
/ PRIOR FILING DATE: 2002-05-08
/ Prior Application removed - See Palm or File Wrapper
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 10
/ LENGTH: 594
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-141-761-10

```

Query Match 5.5%; Score 52.4; DB 15; Length 594;
Best Local Similarity 7.0%; Pred. No. 5,2e-06;

Matches 34; Conservative 189; Mismatches 265; Indels 0; Gaps 0;

```

QY 393 ATACATAACAAGACACATCTCACTACAGAGAGATGATGATCACCAGTCCGTCGGCGG 452
DB 60 AEPVNMWPMQGDITNMIDRFVRAHLDHIDPYPLLTITTSPEQESDERKCNVERXRL 119
QY 453 CGAATGCCACTTCGCTCGTCGTCGTCATGATGATCACTCAACACACCGGTGCGG 512
DB 120 VQNDPAGISEBQCLYQIYIDELYGQLQRPSEDEKXLAEKASIGTYEDSVAVEKAA 179
QY 513 CGCACANTATGATGATGATTTGCGCGGATCAACAGCATGATTTCCCGCGTGGCT 572
DB 180 EKPEEESAAEESNSDEVDIVPIDVVDVDELNQVADLNKQATTYGMADGDFVRL 239
QY 573 GATGAAATATGCGGTACCGCGCATCAATTCGCCGGAGCTACAGACTTAAAGTTG 632
DB 240 RKDKEEABAIKHALEBEKAMYSGRSRRRQREFREKRLRGKTISSPYARSDPTYP 299
QY 633 ACCCGCGAGCGCGGAGAGAGAGAGCTTTCGACGGAATGCCAAACAGAGAGCGCGCAGAC 692
DB 300 YKRSPESSSSSRSSRSPTPGREGKITFTTSFGSDSEAAAAAAGVTTGKPPAP 359
QY 693 CGCGAGCGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 752
DB 360 PQGGPAPGRNARSARRSSSSSSSSASRTSSSSSSSSSSSSSSSSSSSSSSSSSSSS 419
QY 753 GAGACAGCGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 812
DB 420 RSMWSRSTRSRYSRSGRGRHSGGSRDGHYSRSPARRGGYPRRRSRSSHSQDRY 479
QY 813 CCGAAGCAGCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 872
DB 480 RRGGRGLHHSSSRSSWSLSPSRSLTRSRSHSPSPQSRSRSSRSGSPSPARE 539
QY 873 CCGGCCAC 880
DB 540 KLTRPAA 547

```

Db	540	KLTRPAS	547
Query Match	Best Local Similarity	5.5%; Score 52.4; DB 15; Length 594;	
Matches	34; Conservative	189; Mismatches 265; Indels 0; Gaps 0	
US-10-158-790-10			
RESULT 10			
US-10-158-790-10			
Sequence 10, Application US/10158790			
Publication No. US20030180879A1			
GENERAL INFORMATION:			
APPLICANT: Baker, Kevin P.			
APPLICANT: Bereslin, Maureen			
APPLICANT: DeForge, Laura			
APPLICANT: Desnoyers, Luc			
APPLICANT: Filvarsoff, Ellen			
APPLICANT: Gao, Wei-Qiang			
APPLICANT: Gerritsen, Mary E.			
APPLICANT: Goddard, Audrey			
APPLICANT: Godowski, Paul J.			
APPLICANT: Gurney, Austin L.			
APPLICANT: Sherwood, Steven			
APPLICANT: Smith, Victoria			
APPLICANT: Stewart, Timothy A.			
APPLICANT: Tumas, Daniel			
APPLICANT: Watanabe, Colin K			
APPLICANT: Wood, William			
APPLICANT: Zhang, Zemin			
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			
FILE REFERENCE: P3330R1C448			
CURRENT APPLICATION NUMBER: US/10/158,790			
CURRENT FILING DATE: 2002-05-30			
Prior Application removed - See file wrapper or Palm			
NUMBER OF SEQ ID NOS: 550			
SEQ ID NO 10			
LENGTH: 594			
TYPE: PRT			
ORGANISM: Homo Sapien			
US-10-158-790-10			
Query Match	Best Local Similarity	5.5%; Score 52.4; DB 15; Length 594;	
Matches	34; Conservative	189; Mismatches 265; Indels 0; Gaps 0	
US-10-158-790-10			
RESULT 10			
US-10-158-790-10			
Sequence 10, Application US/10158790			
Publication No. US20030180879A1			
GENERAL INFORMATION:			
APPLICANT: Baker, Kevin P.			
APPLICANT: Bereslin, Maureen			
APPLICANT: DeForge, Laura			
APPLICANT: Desnoyers, Luc			
APPLICANT: Filvarsoff, Ellen			
APPLICANT: Gao, Wei-Qiang			
APPLICANT: Gerritsen, Mary E.			
APPLICANT: Goddard, Audrey			
APPLICANT: Godowski, Paul J.			
APPLICANT: Gurney, Austin L.			
APPLICANT: Sherwood, Steven			
APPLICANT: Smith, Victoria			
APPLICANT: Stewart, Timothy A.			
APPLICANT: Tumas, Daniel			
APPLICANT: Watanabe, Colin K			
APPLICANT: Wood, William			
APPLICANT: Zhang, Zemin			
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			
FILE REFERENCE: P3330R1C448			
CURRENT APPLICATION NUMBER: US/10/158,790			
CURRENT FILING DATE: 2002-05-30			
Prior Application removed - See file wrapper or Palm			
NUMBER OF SEQ ID NOS: 550			
SEQ ID NO 10			
LENGTH: 594			
TYPE: PRT			
ORGANISM: Homo Sapien			
US-10-158-790-10			

Qy	873	CCGGCCAC	880
		⋮	⋮
Db	540	KLTRPAAS	547

```

RESULT 11
US-10-137-871-10
: Sequence 10, Application US/10137871
: Publication No. US20030207350A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Guiney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin X
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P330R1C153
CURRENT APPLICATION NUMBER: US/10/137,871
CURRENT FILING DATE: 2002-05-03
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 10
LENGTH: 594
TYPE: PRT
ORGANISM: Homo Sapien
US-10-137-871-10

```

Qy	813	CCGACACAGCGGCCAAGGCACCGCGCAAGCACAACGGCTTAAAGCCCTGCGCCACGG	8722
Dd	480	RGGGRLRHSSSRSSWLSLPSMSRLTRSHSPSPGSRGRSRSGSPSPARE	529
Qy	873	CCGGCCAC	880
Dd	540	KLTRPALS	547

```

RESULT 12
US-10-140-923-10
; Sequence 10, Application US/10140923
; Publication No. US20030207355A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Geritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gueney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C18
CURRENT APPLICATION NUMBER: US/10/140,923
CURRENT FILING DATE: 2002-05-07
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 10
LENGTH: 594
TYPES: PRT
ORGANISM: Homo Sapien
US-10-140-923-10

```



```
QY 693 CGGACGAGACCCCGGCGAGAGCCCACTGCGGACGAGCGGCGGACGACAGATAC 752
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 360 POPGPAPGRNMSARRSSSSSSSSASRTSSRSRSSRSRRGGYRSGHMSRS 419
QY 753 GAGACGCGGACGACACCCCGCGTCATGACGACGAGACGCGCCGAGGAGCGGAC 812
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 420 RWSRERSRYSRSGRHSRSGSRDGRYSRSPARRGYGPRRSRSRSHSGDRY 479
QY 813 CGGACGACCGCGGACGACGCGGACGACGACGCGCCTAGGCGCTGCGGACG 872
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 480 RRGGRGLRHSSRSRSSLSPSRSLTRSRSHSPSPGSRSRSRSGSPSPARE 539
QY 873 CCGGCCAC 880
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 540 KLTRPABS 547
```

RESULT 15

```
US-10-140-805-10
/ Sequence 10, Application US/10140805
/ Publication No. US20030207417A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tamas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zhen
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P333ORC176
/ CURRENT APPLICATION NUMBER: US/10/140,805
/ PRIORITY FILING DATE: 2002-05-07
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 10
/ LENGTH: 594
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-140-805-10
```

Query Match 5.5%, Score 52.4, DB 16, Length 594;
Best Local Similarity 7.0%, Pred. No. 5, 2e-06;
Matches 34, Conservative 189, Mismatches 265, Indels 0, Gaps 0;

```
QY 393 ATACATAACAAGACACTCACTACGAGTGCATGATCAGCGGTCGCGGCGG 452
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 AESPNNMPPWOGDNNMNDRFVRALHLPDYPPLITTSPPQESDERKNTYRGL 119
QY 453 CGAATGCCACTTGGCTCTGCTGCGCATGATGATGATGATGATGATGATGATG 512
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 VQNDPAGISEBQCLYQIYIDELVGLQPSSEDEKKLAEKASIGTYEDSTVAEVEKA 179
QY 513 CGCACANTATAGAGTATTCGCGCGGATCAAGACATGATTTCCCGCGTGGCT 572
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 EKPEEESAABEESNSDEVIPDIDVEVDDELNOQVADLNMKQATTYGMADGDFVRL 239
QY 573 GATCGAATATGCGGACGCGGATCAATTCGCGGAGCTACAGACCTAAAAAGTGTG 632
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 RKDKKEAFAIKHAKALEEKAMYSGRSRKRRREFREKRLGRKISPPSTARSDSPYDP 299
```

```
QY 633 ACCGCGACGCGCGGACGACGCTTTGACGAAATGCCAAACACAGAGGCGCGAC 692
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 300 YKSPSESSSESRSRSPTPGRBKXITFTITSGSDBEFAAAAAAAAGVTTGKPPAP 359
QY 693 CGGACGCGACCCCGGCGGAGAGCCCACTGCGGACGAGCGGCGGACGACAGATAC 752
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 360 POPGPAPGRNMSARRSSSSSSASRTSSRSRSSRSRRGGYRSGHMSRS 419
QY 753 GAGACGCGGACGACACCCCGCGTCATGACGACGAGACGCGCCGAGGAGCGGAC 812
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 420 RWSRERSRYSRSGRHSRSGSRDGRYSRSPARRGYGPRRSRSRSHSGDRY 479
QY 813 CGGACGACCGCGGACGACGCGGACGACGACGCGCCTAGGCGCTGCGGACG 872
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 480 RRGGRGLRHSSRSRSSLSPSRSLTRSRSHSPSPGSRSRSRSGSPSPARE 539
QY 873 CCGGCCAC 880
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 540 KLTRPABS 547
```

Search completed: April 26, 2004, 22:54:08
Job time : 405.717 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 15:46:02 ; Search time 2327.37 seconds
(without alignments)
12279.158 Million cell updates/sec

Title: US-10-076-747-7
Perfect score: 957
Sequence: 1 ttgggcttactaatgcatg.....ggcagcagagaccacgcagct 957

Scoring table:
IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 1491090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_gestba:*
2: em_gesthum:*
3: em_estin:*
4: em_estum:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hnc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hnc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estum:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vit:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	81.6	8.5	1018	12	BM914338 AGENCOURT
2	69.6	7.3	891	14	CD107024 AGENCOURT
3	67.4	7.0	770	14	BG686134 602638481
4	66.6	7.0	1201	13	BX356664 BX356664

5	66.4	6.9	1201	9	AL531683	AL531683
6	65.4	6.8 <td>753</td> <td>13 <td>BU933171 <td>BU933171 AGENCOURT</td> </td></td>	753	13 <td>BU933171 <td>BU933171 AGENCOURT</td> </td>	BU933171 <td>BU933171 AGENCOURT</td>	BU933171 AGENCOURT
7	65.4	6.8 <td>939</td> <td>29 <td>CNS004NE <td>AL054280 Drosophila</td> </td></td>	939	29 <td>CNS004NE <td>AL054280 Drosophila</td> </td>	CNS004NE <td>AL054280 Drosophila</td>	AL054280 Drosophila
8	64.8	6.8 <td>925</td> <td>29 <td>CNS0081P <td>AL053013 Drosophila</td> </td></td>	925	29 <td>CNS0081P <td>AL053013 Drosophila</td> </td>	CNS0081P <td>AL053013 Drosophila</td>	AL053013 Drosophila
9	64.6	6.8 <td>1144</td> <td>13 <td>BX415926 <td>BX415926 BX415926</td> </td></td>	1144	13 <td>BX415926 <td>BX415926 BX415926</td> </td>	BX415926 <td>BX415926 BX415926</td>	BX415926 BX415926
10	62.6	6.5 <td>1100</td> <td>29 <td>CNS016KD <td>AL106856 Drosophila</td> </td></td>	1100	29 <td>CNS016KD <td>AL106856 Drosophila</td> </td>	CNS016KD <td>AL106856 Drosophila</td>	AL106856 Drosophila
11	62.6	6.5 <td>1213</td> <td>12 <td>BM918618 <td>BM918618 AGENCOURT</td> </td></td>	1213	12 <td>BM918618 <td>BM918618 AGENCOURT</td> </td>	BM918618 <td>BM918618 AGENCOURT</td>	BM918618 AGENCOURT
12	62.4	6.5 <td>1101</td> <td>29 <td>CNS017SY <td>AL108460 Drosophila</td> </td></td>	1101	29 <td>CNS017SY <td>AL108460 Drosophila</td> </td>	CNS017SY <td>AL108460 Drosophila</td>	AL108460 Drosophila
13	61.4	6.4 <td>844</td> <td>29 <td>CNS0052P <td>AL056552 Drosophila</td> </td></td>	844	29 <td>CNS0052P <td>AL056552 Drosophila</td> </td>	CNS0052P <td>AL056552 Drosophila</td>	AL056552 Drosophila
14	61.2	6.4 <td>515</td> <td>13 <td>BX424977 <td>BX424977 BX424977</td> </td></td>	515	13 <td>BX424977 <td>BX424977 BX424977</td> </td>	BX424977 <td>BX424977 BX424977</td>	BX424977 BX424977
15	61	6.4 <td>510</td> <td>29 <td>CNS006ON <td>AL065629 Drosophila</td> </td></td>	510	29 <td>CNS006ON <td>AL065629 Drosophila</td> </td>	CNS006ON <td>AL065629 Drosophila</td>	AL065629 Drosophila
16	60.8	6.4 <td>958</td> <td>13 <td>BU146944 <td>BU146944 AGENCOURT</td> </td></td>	958	13 <td>BU146944 <td>BU146944 AGENCOURT</td> </td>	BU146944 <td>BU146944 AGENCOURT</td>	BU146944 AGENCOURT
17	60.2	6.3 <td>1201</td> <td>13 <td>BX340009 <td>BX340009 BX340009</td> </td></td>	1201	13 <td>BX340009 <td>BX340009 BX340009</td> </td>	BX340009 <td>BX340009 BX340009</td>	BX340009 BX340009
18	60	6.3 <td>1067</td> <td>10 <td>BX457099 <td>BX457099 BX457099</td> </td></td>	1067	10 <td>BX457099 <td>BX457099 BX457099</td> </td>	BX457099 <td>BX457099 BX457099</td>	BX457099 BX457099
19	60	6.3 <td>1160</td> <td>13 <td>BF129273 <td>BF129273 60.810871</td> </td></td>	1160	13 <td>BF129273 <td>BF129273 60.810871</td> </td>	BF129273 <td>BF129273 60.810871</td>	BF129273 60.810871
20	59.8	6.2 <td>1103</td> <td>13 <td>BX403654 <td>AL063654 Drosophila</td> </td></td>	1103	13 <td>BX403654 <td>AL063654 Drosophila</td> </td>	BX403654 <td>AL063654 Drosophila</td>	AL063654 Drosophila
21	59.2	6.2 <td>932</td> <td>29 <td>CNS0072Q <td>AL065742 Drosophila</td> </td></td>	932	29 <td>CNS0072Q <td>AL065742 Drosophila</td> </td>	CNS0072Q <td>AL065742 Drosophila</td>	AL065742 Drosophila
22	59	6.2 <td>624</td> <td>13 <td>BX404419 <td>BX404419 BX404419</td> </td></td>	624	13 <td>BX404419 <td>BX404419 BX404419</td> </td>	BX404419 <td>BX404419 BX404419</td>	BX404419 BX404419
23	58.8	6.1 <td>765</td> <td>13 <td>BUS95659 <td>BUS95659 AGENCOURT</td> </td></td>	765	13 <td>BUS95659 <td>BUS95659 AGENCOURT</td> </td>	BUS95659 <td>BUS95659 AGENCOURT</td>	BUS95659 AGENCOURT
24	58.8	6.1 <td>1201</td> <td>13 <td>BX360624 <td>BX360624 BX360624</td> </td></td>	1201	13 <td>BX360624 <td>BX360624 BX360624</td> </td>	BX360624 <td>BX360624 BX360624</td>	BX360624 BX360624
25	58.6	6.1 <td>932</td> <td>29 <td>CNS0072Q <td>AL065742 Drosophila</td> </td></td>	932	29 <td>CNS0072Q <td>AL065742 Drosophila</td> </td>	CNS0072Q <td>AL065742 Drosophila</td>	AL065742 Drosophila
26	58.6	6.1 <td>939</td> <td>13 <td>BQ706586 <td>BQ706586 AGENCOURT</td> </td></td>	939	13 <td>BQ706586 <td>BQ706586 AGENCOURT</td> </td>	BQ706586 <td>BQ706586 AGENCOURT</td>	BQ706586 AGENCOURT
27	58.2	6.1 <td>937</td> <td>29 <td>CNS006XP <td>AL066056 Drosophila</td> </td></td>	937	29 <td>CNS006XP <td>AL066056 Drosophila</td> </td>	CNS006XP <td>AL066056 Drosophila</td>	AL066056 Drosophila
28	58.2	6.1 <td>1095</td> <td>9 <td>AL551270 <td>AL551270 AL551270</td> </td></td>	1095	9 <td>AL551270 <td>AL551270 AL551270</td> </td>	AL551270 <td>AL551270 AL551270</td>	AL551270 AL551270
29	58	6.1 <td>912</td> <td>29 <td>CNS006N3 <td>AL065775 Drosophila</td> </td></td>	912	29 <td>CNS006N3 <td>AL065775 Drosophila</td> </td>	CNS006N3 <td>AL065775 Drosophila</td>	AL065775 Drosophila
30	57.6	6.0 <td>886</td> <td>13 <td>BQ712155 <td>BQ712155 AGENCOURT</td> </td></td>	886	13 <td>BQ712155 <td>BQ712155 AGENCOURT</td> </td>	BQ712155 <td>BQ712155 AGENCOURT</td>	BQ712155 AGENCOURT
31	57.6	6.0 <td>924</td> <td>13 <td>BX442207 <td>BX442207 BX442207</td> </td></td>	924	13 <td>BX442207 <td>BX442207 BX442207</td> </td>	BX442207 <td>BX442207 BX442207</td>	BX442207 BX442207
32	57.2	6.0 <td>1200</td> <td>13 <td>BX415896 <td>BX415896 BX415896</td> </td></td>	1200	13 <td>BX415896 <td>BX415896 BX415896</td> </td>	BX415896 <td>BX415896 BX415896</td>	BX415896 BX415896
33	57	6.0 <td>1201</td> <td>29 <td>CNS016BR <td>AL106545 Drosophila</td> </td></td>	1201	29 <td>CNS016BR <td>AL106545 Drosophila</td> </td>	CNS016BR <td>AL106545 Drosophila</td>	AL106545 Drosophila
34	56.8	5.9 <td>932</td> <td>29 <td>CNS0070E <td>AL065254 Drosophila</td> </td></td>	932	29 <td>CNS0070E <td>AL065254 Drosophila</td> </td>	CNS0070E <td>AL065254 Drosophila</td>	AL065254 Drosophila
35	56.6	5.9 <td>935</td> <td>29 <td>CNS006KX <td>AL065051 Drosophila</td> </td></td>	935	29 <td>CNS006KX <td>AL065051 Drosophila</td> </td>	CNS006KX <td>AL065051 Drosophila</td>	AL065051 Drosophila
36	56.4	5.9 <td>1159</td> <td>29 <td>CNS015XR <td>AL106041 Drosophila</td> </td></td>	1159	29 <td>CNS015XR <td>AL106041 Drosophila</td> </td>	CNS015XR <td>AL106041 Drosophila</td>	AL106041 Drosophila
37	56.2	5.9 <td>784</td> <td>29 <td>AG161654 <td>AG161654 Pan trogl</td> </td></td>	784	29 <td>AG161654 <td>AG161654 Pan trogl</td> </td>	AG161654 <td>AG161654 Pan trogl</td>	AG161654 Pan trogl
38	56	5.9 <td>938</td> <td>29 <td>CNS00KM6 <td>AL0707749 Drosophila</td> </td></td>	938	29 <td>CNS00KM6 <td>AL0707749 Drosophila</td> </td>	CNS00KM6 <td>AL0707749 Drosophila</td>	AL0707749 Drosophila
39	56	5.9	1015	28 <td>BZ569259 <td>BZ569259 pac82-164</td> </td>	BZ569259 <td>BZ569259 pac82-164</td>	BZ569259 pac82-164
40	55.8	5.8 <td>1165</td> <td>13 <td>BX425080 <td>BX425080 BX425080</td> </td></td>	1165	13 <td>BX425080 <td>BX425080 BX425080</td> </td>	BX425080 <td>BX425080 BX425080</td>	BX425080 BX425080
41	55.8	5.8 <td>1201</td> <td>13 <td>BX405071 <td>BX405071 BX405071</td> </td></td>	1201	13 <td>BX405071 <td>BX405071 BX405071</td> </td>	BX405071 <td>BX405071 BX405071</td>	BX405071 BX405071
42	55.4	5.8 <td>1101</td> <td>29 <td>CNS00BNG <td>AL057398 Drosophila</td> </td></td>	1101	29 <td>CNS00BNG <td>AL057398 Drosophila</td> </td>	CNS00BNG <td>AL057398 Drosophila</td>	AL057398 Drosophila
43	55.2	5.8 <td>894</td> <td>29 <td>CNS01591 <td>AL105168 Drosophila</td> </td></td>	894	29 <td>CNS01591 <td>AL105168 Drosophila</td> </td>	CNS01591 <td>AL105168 Drosophila</td>	AL105168 Drosophila
44	55.2	5.8 <td>953</td> <td>13 <td>BQ709520 <td>BQ709520 AGENCOURT</td> </td></td>	953	13 <td>BQ709520 <td>BQ709520 AGENCOURT</td> </td>	BQ709520 <td>BQ709520 AGENCOURT</td>	BQ709520 AGENCOURT
45	55.2	5.8 <td>1200</td> <td>29 <td>CNS01671 <td>AL106392 Drosophila</td> </td></td>	1200	29 <td>CNS01671 <td>AL106392 Drosophila</td> </td>	CNS01671 <td>AL106392 Drosophila</td>	AL106392 Drosophila

ALIGNMENTS

RESULT 1
BM914338 1018 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT 6615480 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:5480006
DEFINITION 5', mRNA sequence.
ACCESSION BM914338
VERSION BM914338.1 GI:19364717
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1018)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strusberg, Ph.D.
Email: cgsbds-remail.nih.gov
Tissue Procurement: Dr. Mark Watson
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2002 row: C column: 15
High quality sequence stop: 644.

FEATURES

source

Location/Qualifiers

1. 1018
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5480006"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_113"
 /note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 8.5%; Score 81.6; DB 12; Length 1018;
 Best Local Similarity 62.6%; Pred. No. 5.9e-08;
 Matches 218; Conservative 0; Mismatches 109; Indels 21; Gaps 5;

QY 63 AGTACGGGGTTCAGAGCTATCTGAGCTGACGGCTGATGCACTGTGAGAGTCCACAG 122
 DB 591 AAGTACGGGGTTCAGAGCTATCTGAGCTGACGGCTGATGCACTGTGAGAGTCCACAG 646
 QY 123 AAGTACGGGGTTCAGAGCTATCTGAGCTGACGGCTGATGCACTGTGAGAGTCCACAG 182
 DB 647 AAGTACGGGGTTCAGAGCTATCTGAGCTGACGGCTGATGCACTGTGAGAGTCCACAG 698
 QY 183 GGTCCCTTACGAAATGTTCTATAGTTCTTAAACAGCTGACCGCCCACTACCAATGCC 242
 DB 699 GGGCCCCCTACGAAATGTTCTATAGTTCTTAAACAGCTGACCGCCCACTACCAATGCC 753
 QY 243 GAGACTACAGCTGACGAGATCCCAAGGAGAGAGGGGTCTCTCATCCACACACACAG 302
 DB 754 AAAATTAGAGCTGACGAGATCCCAAGGAGAGAGGGGTCTCTCTCCCTCCCAACCCACAG 813
 QY 303 GCGGATCAAGCCCTTATCATGCGGATGTGACGTCATGTAAGAGCGCTCAAAATTA 362
 DB 814 GGAATTAAAGCCCTTATCATGCGGATGTGACGTCATGTAAGAGCGCTCAAAATTA 869
 QY 363 GATATTCGCACTGTTGAAATGTTCTTCACTACATCAATAAGAGAGCA 410
 DB 870 AAAATTCCTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 917

RESULT 2 891 bp mRNA linear EST 15-MAY-2003
 CD107024
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

CD107024 891 bp mRNA linear EST 15-MAY-2003
 AGNCOURT 14020401 NIH_MGC_179 Homo sapiens cDNA clone
 IMAGE:30368459 5', mRNA sequence.
 CD107024
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Invitrogen Corp
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: NDAM431 row: e column: 12

FEATURES

source

High quality sequence stop: 533.

Location/Qualifiers

1. 891
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30368459"
 /tissue_type="Pituitary"
 /lab_host="DH10B-Tom A (T1 and T5 phage resistancees)"
 /clone_lib="NIH_MGC_179"
 /note="Organ: brain; Vector: pCMV-Sport6.1; Site_1: EcoRV (destroyed); Site_2: NotI; Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.1 kb. Library was constructed by (Invitrogen). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 7.3%; Score 69.6; DB 14; Length 891;
 Best Local Similarity 64.1%; Pred. No. 3.1e-05;
 Matches 159; Conservative 0; Mismatches 74; Indels 15; Gaps 3;

QY 63 AGTACGGGGTTCAGAGCTATCTGAGCTGACGGCTGATGCACTGTGAGAGTCCACAG 122
 DB 594 AAGTACGGGGTTCAGAGCTATCTGAGCTGACGGCTGATGCACTGTGAGAGTCCACAG 649
 QY 123 AAGTACGGGGTTCAGAGCTATCTGAGCTGACGGCTGATGCACTGTGAGAGTCCACAG 182
 DB 650 AAGTACGGGGTTCAGAGCTATCTGAGCTGACGGCTGATGCACTGTGAGAGTCCACAG 703
 QY 183 GGTCCCTTACGAAATGTTCTATAGTTCTTAAACAGCTGACCGCCCACTACCAATGCC 242
 DB 704 GGGCCCCCTACGAAATGTTCTATAGTTCTTAAACAGCTGACCGCCCACTACCAATGCC 758
 QY 243 GAGACTACAGCTGACGAGATCCCAAGGAGAGAGGGGTCTCTCATCCACACACACAG 302
 DB 759 GAAAGCTAAATCTGACAGGATCCCAAGGAGAGGGGTCTCTCTCCCAACCCCAAGGC 818
 QY 303 GCGGATC 310
 DB 819 ATCAAGTC 826

RESULT 3 770 bp mRNA linear EST 01-MAY-2001
 BG686194/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

BG686194 770 bp mRNA linear EST 01-MAY-2001
 60263848F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4766108 5',
 mRNA sequence.
 BG686194
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LILCM1626 row: a column: 21
 High quality sequence stop: 307.

FEATURES

source

Location/Qualifiers

1. 770
 /organism="Homo sapiens"

```

/mol type="mrna"
/db xref="taxon:9606"
/clone="IMAGE:4766108"
/tissue type="primary B-cells from tonsils (cell line)"
/lab host="DH10B (phage-resistant)"
/clone.lib="NIH MGC 48"
/note="Organ: B-cells; Vector: pOTB7; Site:1: XhoI;
Site:2: EcoRI. cDNA made by oligo-dT priming;
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC library."

```

ORIGIN		
Query Match	7.0%;	Score 67.4; DB 12; Length 770;
Best Local Similarity	61.6%;	Pred. No. 9.3e-05;
Matches 141; Conservative	0;	Mismatches 86; Indels 2; Gaps 2

QY	68	CGGGGTCAGAGGCTATCTAGAGCTGACGGGCTCGATGAGTGTAGAGATCCACAGAAAGCT	127
	333	CGGGCCGACAGACTTACTTAAAGCTCGACGCTTGACACGTGAGAGATGCCACCGAAAGCT	274
Db			
QY	128	ACAAGCTGCACAGAGTCTTAGTGTGATGATAGGAGACACCGTGGATGAGACAGTGGTCC	187
	273	ACAAGCTGC-CACGGTCAACGCATCGAAGAAAGGAGGACCGGTGGAGAAAGACATAGGCC	215
Db			
QY	188	CTTACAGAAATGTTCTATAGTGTCTCTTAAACAGCTCAGCCCCCTTACCAATGGCGAGAC	247
	214	CTTACAGAAATGTTCTATAGTGTCTCTTCAACCCCTCAACCCCCACACCAACGGAGGAAAC	155
Db			
QY	248	TAGCAGGTGACGAGGATGCCAAGGAGAGGGGTCTCTTCATCTCACACCA	296
	154	TAAAGCTGCAC-GGAGTTCACAGGGAGAGGGGTCTTCTTCTCCACCCCA	107
Db			

RESULT 4	
BX356664/c	
LOCUS	
DEFINITION	BX356664 1201 bp RNA linear EST 05-MAY-2001
ACCESSION	EX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
VERSION	Clone CS0D015YB03 3-PRIME, mRNA sequence.
KEYWORDS	BX356664
SOURCE	BX356664.1 GI:30378083
ORGANISM	EST.
COMMENT	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
CONTRACT	Full-length cDNA libraries and normalization
	Unpublished (2001)
	Contact: Genoscope
	Genoscope - Centre National de Sequencage
	BP 191 91006 EVRY cedex - France
	Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
	library was constructed by Life Technologies, a division of
	Invitrogen Contact : Feng Liang Email : fliang@life-tech.com URL :
	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
	Faraday Avenue Genoscope sequence ID : CS0D015CA02NPF1.

```

FEATURES
    source
        Location/Qualifiers
            1..1201
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="raxon.9606"
                /clone="CSOD.015XP03"
                /cdate_type="PLACENTA COT 25-NORMALIZED"
                /clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
                /note="1st strand cDNA was primed with a NotI-oligo(dT)
                    primer. Five prime end enriched, double-strand cDNA was
                    digested with NotI and cloned into the NotI and EcoR V

```

```

ORIGIN                                     sites of the pCMVSPORT 6 vector. Library was normalized.
Query Match                               7.0%; Score 66.6; DB 13; Length 1201;
Best Local Similarity 11.5%; Pred. No. 0.00017;
Matches 48; Conservative 215; Mismatches 154; Indels 0; Gaps 0

```

QY	538	CCGGGACACAGACCATGAGATTTCCTCCCCCGGCGTGATCGAATATGCGCGTACGGCGGATC	597
Db	1071	VSVAVAAAAAVVSSSSSSSSSSSSSVSSSSSSSSSMNMSSSSAVASSSSSSSSSSANSSVSS	1011
QY	598	AAATTCCGCCGGAGCTACAGACCTAAAAAAATTGACCGCGCAGCGCGCGGAGACAAGAC	657
Db	1011	SVASSSAAAAASVVAAVVSAAASVSABAAAAAAAVVAASSSSSSSSVSSSSXSAVSSSSS	952
QY	658	TTTGACGCGAATGCCAAACACAGAGGGCCGACACCGCGCAGGCGGACCCCGGGCGGAGAG	717
Db	951	VSSSSSNAVSSSSSSSAAASVSSSAAAVSSVVAASSSSSSSSVAAASSSSAAAAVAAAV	892
QY	718	CCCCACGCGGACAGGCGGCGGCGAAAGACAGATCGAGAGCGCGACCGCGCGCGCGCC	777
Db	891	SAAAAVVSWSVVAASASASVASVSVSSSSSAAANASSASASVAAAAASSSSSVAVVAAAA	832
QY	778	GTGCATGAGACGAGACGCGCCGAGGAGACGCGACCCGAAACAGCGCCGCAAGACGCAAC	837
Db	831	VSSWASAAASAAVSSSSSSSAGAVSSASXASVASSASVBSAGSGSSAAVSVSRESS	772
QY	838	GCGGACCGACACGCGCGCTTAGCGCCCTCGGCCACGCGCGGCGCAACGCGCGAGAGCGGGCG	897
Db	771	VAVVAASVVAAGVAVVAASASASVASSSVMAAAASNAASVVASSSAVAAAAASSSV	712
QY	898	GAGCACCGCAGGAACCGACCAACGAGCCGACCGCGCCAGGCGACGAGCGACCACGA	954
Db	711	SAAASASSSAAASSSASSSSAAAVVSSSSSSSSSSSVASVAVVSSASSSASSSSA	655

FEATURES	RESULT 5	LOCUS	AL531683	DEFINITION	AL531683 Homo sapiens FETAL LIVER Homo sapiens cDNA clone	1201 bp	linear	EST 23-MAY-2003
ACCESSION	AL531683	AL531683	AL531683	CSODM0081T01 3-PRIME, mRNA sequence.				
VERSION	AL531683.2	AL531683.2	AL531683.2	GI:31069515				
KEYWORDS	EST.							
SOURCE	Homo sapiens (human)							
ORGANISM	Homo sapiens							
REFERENCE	Eukaryotic; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.							
TITLE	1 (bases 1 to 1201)							
JOURNAL	Li,W.B., Gruber,C., Jesse,J., and Polayes,D.							
COMMENT	Full-length cDNA libraries and normalization							
	Unpublished (2001)							
	On Feb 13, 2001 this sequence version replaced gi:112795176.							
	Contact: Genoscope							
	Genoscope - Centre National de Sequencage							
	BP 191 91067 EVRY cedex - France							
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr							
	Library was constructed by Life Technologies, a division of							
	Invitrogen. Contact : Feng Liang Email : fliang@lifestech.com URL :							
	http://Fulllength.invitrogen.com/Invitrocdn Corporation 1600							
	Faraday Avenue Genoscope sequence ID : CSODM008AE01NP1.							
	Location/Qualifiers							

```

source
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GSDMD008Y101"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_id="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-cligo (dt) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and

```

cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

Query Match	6.9%;	Score 66.4;	DB 9;	Length 1201;
Best Local Similarity	33.6%;	Pred. No. 0.00019;		
Matches 103;	Conservative 73;	Mismatches 131;	Indels 0;	Gaps 0;

[illegible]

RESULT 6
LOCUS BU933171
DEFINITION AGN002T.1047293; NIH_MGC_127 Homo sapiens CDNA clone
IMAGE:6673502 5', mRNA sequence.
BU933171
ACCESSION BU933171.1 GI:24121990
VERSION EST.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 753)
NIH-MGC <http://mgc.ni.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

cDNA Library Preparation: Michael Brownstein Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MCC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINTL at:
<http://image.llnl.gov>
 Plate: LINC2956 row: m column: 06
 High quality sequence: 197.

```

1..753
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6673902"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phase-resistant)"
/clone_id="NH MGC 127"
/notes="Vector: pNCR-11B. Site 1: SfiI (ggccatcgtgccc);
Site 2: SfiI (ggccatcgtgccc); Double-stranded cDNA was
prepared from a pool of 40 cell line poly(A+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -

```

Query Match	6.8%;	Score 65.4;	DB 13;	Length 753;
Best Local Similarity	58.0%;	Pred. No. 0.0027;		
Matches 192;	Conservative 0;	Mismatches 131;	Indels 8;	Gaps 4

4th, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:
5'-AAGCAGAGTGTATACGACGAGATGGCTTACGGCCGGG-3' and
5'-ATTTCAGAGCCGACGACGCGACATC-dt(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 1-2 kb size fraction (other fractions present in NIH_MGC_126 and NIH_MGC_128). Library created in the laboratory of T. Udell, M.D., Ph.D. (NIH). Note: this is a NIH_MGC Library."

Qy	7	CAGCAGCTCTCTGAG--CTGCACGCGCTATGCACTGTGACAGTCC--CACAGAACTAC	129
Db	241	CAGCAGCTCTCTGAGCCCTGAGCCCTGAGCACTGTGAGAAAGTCCCAAGAAAGCTAC	300
Qy	130	AGCCTGCAAGAGTCTTAGCTGCATGAAGGAGACACCGTGTGATGAAGACATGTGTCCC	189
Db	301	CGCCTGCGCAGGGTCCCGCATGAAGGAGACACCGTGTGAAAAAACAATGTGTGCTCC	360
Qy	190	T--ACAGAAATTTCTATAGTTCTCTAACGCTCAAGCCCCCACTACCATGTGCGAGAC	247
Db	361	CTAACAGAAATGTTTCATATGAGTTCTTAAACCCCTCAACCCCCCAACAGGGAGAAC	420
Qy	248	TACACGCTG--CAGGATCCCAAGGAGAGGGGTCTCCATCCACACACACAGGCGC	305
Db	421	TAAAAAGCTGGCAGGATCCCAAGGGGAAAGGGGTTTTTTTTCTTCCACCCCAAG	480
Qy	306	GAGTCAAGCCCTTATCATATCGGCATGTGCAAGTCATGTAAAGCGCCTCAAAATAAGT	365
Db	481	GCATTTCAAGCCCCCTTCTCCCTCGGAGCTTTCATTAATAAACCCCTCAATTAATAAAA	540
Qy	366	AATCTGCAGTTGTGAATGTCTACACTAC	396
Db	541	AATTCCTTTGGGCCCAATCCCGAAAAAC	571

RESULT 7					
CNS004NB					
LOCUS					
DEFINITION	CNS004NB	839 bp	DNA	linear	GSS 03-JUN-1999
	Drosophila melanogaster genome survey sequence TET3 end of BAC #				
	BACR0036 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL054280				
VERSION	AL054280.1				
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 839)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
COMMENT	- Web : www.genoscope.cns.fr) determination of this BAC-end sequence was carried out as part of a project funded by the French Ministry of Research.				

Determination of this BAC and sequence was carried out as part of collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org>. The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Oseegawa and Aaron Mammossier in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial

FEATURES
source
Faraday Avenue Genoscope sequence ID : CS0CAP008CD010P1.
Location/Qualifiers
1.1144

/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="CS0CAP008CD010P1"
/feature_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="Vector: PCWVSPT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the PCWVSPT 6 vector. Library was not normalized."

ORIGIN

Query Match 6.8%; Score 64.6; DB 13; Length 1144;
Best Local Similarity 35.9%; Pred. No. 0.0047;
Matches 116; Conservative 61; Mismatches 146; Indels 0; Gaps 0;

QY 626 AAAGTTGACCGCGCGACCGCCGGAAGAACAGCTTTCAGCGAATCCAAACACAGAGGAC 685
DB 751 AAVGCGCGCGSSSSGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 810
QY 686 CCGACAGCCCGCAGCG 745
DB 811 SGC CG 870
QY 746 CAGATTCGAGACCG 805
DB 871 CCG 930
QY 806 CCGACAGCCCGCAGCG 865
DB 931 GGGCG 990
QY 866 GCCACG 925
DB 991 SGC CG 1050
QY 926 CGACCG 948
DB 1051 SGCSSSSSSGCG 1073

RESULT 10
CNS016KD 1100 bp DNA linear GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence Sp6 end of BAC
DEFINITION BACN16D22 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION AL106855.1 GI:5624152
VERSION AL106855
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pezomyzeta;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS Genoscope.
TITLE Direct Subcloning
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
Determination of this BAC end and sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CSHL (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
Project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBeIOBAC11.

FEATURES
source
Location/Qualifiers
1.1100

/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN16D22"
/clone_lib="DrosBAC"
/plasmid="pBeIOBAC11"
/note="end : Sp6"

ORIGIN

Query Match 6.5%; Score 62.6; DB 29; Length 1100;
Best Local Similarity 24.8%; Pred. No. 0.0013;
Matches 87; Conservative 143; Mismatches 119; Indels 2; Gaps 2;

QY 589 CGCGCATCAATTGCG 648
DB 1079 MGGM 1020
QY 649 AGAACAGCGCTTTCAGCGAATGCCAAACACAGAGCGCGCGCGCGCGCGCGCGCGCG 708
DB 1019 GCGGAGAGCG 960
QY 709 GCGGAGAGCG 768
DB 959 GCGGAGAGCG 900
QY 769 CG 828
DB 899 MVMGM 841
QY 829 AGACGACCG 888
DB 840 SMVSGCG 781
QY 889 GCGGAGCG 938
DB 780 MGSRSRSGGAGMCAVGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 730

RESULT 11
EM918618 1213 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT_6635161 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5747630
DEFINITION 5', mRNA sequence.
ACCESSION EM918618
VERSION EM918618.1 GI:19369997
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs@remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL12774 row: b column: 15
High quality sequence stop: 537.

FEATURES
source
Location/Qualifiers
1.1213

/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5747630"

ORIGIN

Query Match 6.5%; Score 62.6; DB 12; Length 1213;
 Best Local Similarity 57.1%; Pred. No. 0.0014;
 Matches 214; Conservative 0; Mismatches 149; Indels 12; Gaps 5;

/lab host="DH10B"
 /clone.lib="NIH_MGC_120"
 /note="Organ: pooled pancreas and spleen; Vector:
 pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of spleen and pancreas from 28 yo
 male. Library is oligo-dT primed and directionally cloned
 (EcoRV site is destroyed upon cloning). Average insert
 size 1.5 kb, insert size range 1-2.5 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 025. Note: this is a NIH_MGC Library."

70 CGGTACAGAGCTATCTAGAGCTTACAGGCTGATGAGAGTCCACAGAGCTAC 129
 623 CGGCGACAGCTATCTAGAGCTTACAGGCTGATGAGAGTCCACAGAGCTAC 129
 130 AGCTGACAGAGCTTATGCTGATGAGAGTCCACAGAGTCCACAGAGCTAC 129
 601 CAGCTGACAGAGCTTATGCTGATGAGAGTCCACAGAGTCCACAGAGCTAC 129
 189 CTACAGAGATGCTTATGCTGATGAGAGTCCACAGAGTCCACAGAGCTAC 129
 738 CTACAGAGATGCTTATGCTGATGAGAGTCCACAGAGTCCACAGAGCTAC 129
 249 AGCAGAGCTGAGAGTCCACAGAGTCCACAGAGTCCACAGAGTCCACAGAG 304
 796 AAAAAGCTGAGAGTCCACAGAGTCCACAGAGTCCACAGAGTCCACAGAG 855
 305 CGAGTCAAGAGCTTATGCTGATGAGAGTCCACAGAGTCCACAGAGTCCACAG 364
 856 GATGACAGAGCTTATGCTGATGAGAGTCCACAGAGTCCACAGAGTCCACAG 915
 365 TATCTGACAGTCTGATGAGAGTCCACAGAGTCCACAGAGTCCACAGAGTCC 424
 916 ATTCTCTCAATGAGAGTCCACAGAGTCCACAGAGTCCACAGAGTCCACAG 975
 425 GAGTCAAGAGTCCACAGAGTCCACAGAGTCCACAGAGTCCACAGAGTCCAC 439
 976 TCGGAGTCCACAGAGTCCACAGAGTCCACAGAGTCCACAGAGTCCACAGAG 990

RESULT 12
 CNS0175Y 1101 bp DNA linear GSS 26-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence Sp6 end of BAC
 DEFINITION BAC11P16 of DrosBAC library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 AL108460
 AL108460.1 GI:5628764
 GSS.
 Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 SOURCE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1101)
 Genoscope.
 REFERENCE Direct Submission
 AUTHORS Submitted (23-JUN-1999) Genoscope - Centre National de Sequencage :
 TITLE BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 JOURNAL - Web : www.genoscope.cns.fr)
 COMMENT Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (BDGP) -
 library (Dros BAC) was made by Alain Billand at CEPR (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MGC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector

FEATURES
 source location/Qualifiers
 1..1101
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone.lib="DrosBAC"
 /clone.lib="DrosBAC"
 /plasmid="pBelobAC11"
 /note="end : Sp6"

ORIGIN

Query Match 6.5%; Score 62.4; DB 29; Length 1101;
 Best Local Similarity 17.1%; Pred. No. 0.0015;
 Matches 64; Conservative 178; Mismatches 129; Indels 3; Gaps 1;

583 GCGTACAGAGCTATCTAGAGCTTACAGGCTGATGAGAGTCCACAGAGCTAC 642
 667 GGGGSCCGGKAKGAGVGRVCGAGGASCAVMAADCGGCCAKVACSSSASSSGSSC 726
 643 GGGGAGAGAGAGGCTTTCAGAGGATGCAAAACAGAGGCGCGAGAGCGGAGGA 702
 727 AATSSASRGMVSSACAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGS 786
 703 CCCGAG 762
 787 SCSSASMGVSSGSSCSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGS 846
 763 GCGACAG 822
 847 GAVSSSCRSV 906
 823 GCGGAG 882
 907 VESVVA---SVSSSSSSSSSSSSSVSVSVSVSVSVSVSVSVSVSVSVSVSVSV 963
 883 GCGAG 942
 964 AVMSAVVSV 1023
 943 CAG 956
 1024 AASSSSSSSSSSSVS 1037

RESULT 13
 CNS0052P 844 bp DNA linear GSS 03-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence T8T3 end of BAC #
 DEFINITION BAC11P16 of RPT-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 AL056652
 AL056652.1 GI:4932342
 GSS.
 Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 SOURCE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 844)
 Genoscope.
 REFERENCE Direct Submission
 AUTHORS Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 TITLE BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 JOURNAL - Web : www.genoscope.cns.fr)
 COMMENT Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see http://www.fruitfly.org The BDGP Drosophila
 melanogaster BAC library was prepared by Kazuhiro Oseogawa and
 Aaron Mammose in Peter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo.

NY. The library is named RPCT-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1..844
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BAC11P16"
/clone_lib="RPCT-98"
/note="end : TET3"

ORIGIN

Query Match 6.4%; Score 61.4; DB 29; Length 844;
Best Local Similarity 20.8%; Pred. No. 0.0023;
Matches 45; Conservative 107; Mismatches 64; Indels 0; Gaps 0;

QY 733 GCGAGCGGAGACGATCGAGCGGAGCGGAGCGGCGCGCGCGCGATGAGACGAG 792
Db 369 SSVASSSMAAAGASCBGAAANMCAAGCGSSSSASRSRMSSCSSCASSCGVSSSCSA 428
QY 793 ACGGCGGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 852
Db 423 MSSSCTCVSSCGMASSSCTCGMSASSSSSSSGCVSSVSGRAVARGRCMCMSVCCMC 488
QY 853 GCTTACGCTTCCGCGGAGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 912
Db 489 SSMCMSCSVSSVAVASGSSSVRAVAGCGVGGVSSRGRARSSRGSGSSSVSGV 548
QY 913 CCGACCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 948
Db 549 VSSSSVGMGCGCASSASVSCBSVSSVSSVSSVSSVSSVSSVSSVSSVSSVSSVSS 584

RESULT 14
BX424977 515 bp mRNA linear EST 15-MAY-2003
LOCUS BX424977 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBA005ZH04
DEFINITION 3-PRIME, mRNA sequence.
ACCESSION BX424977.1 GI:30784421
VERSION BX424977
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 515)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT Contact: Genoscope
BP 191 91006 EVRY cedex - France
Email: segefe@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6304.f for more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CLOBA005ZH04&FPLcluster=6304.f>. Contact : Feng Liang Email : liang@lifetech.com URL : Corporation 1600 <http://fulllength.invitrogen.com/>
Faraday Avenue Genoscope sequence ID : CLOBA005ZH04FPL.

FEATURES

source

1..515
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBA005ZH04"
/issue_type="PLACENTA"

/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 6.4%; Score 61.2; DB 13; Length 515;
Best Local Similarity 11.3%; Pred. No. 0.0021;
Matches 27; Conservative 129; Mismatches 83; Indels 0; Gaps 0;

QY 680 GAGCGCGGAGACCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 739
Db 274 GGGGGNAAAAAANTTGGGSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 333
QY 740 GAAGACAGATACGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 799
Db 334 SVVSS 393
QY 800 AGGAGCGGAGACCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 859
Db 394 SSSGSS 453
QY 860 CCTCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 918
Db 454 SGNVSS 512

RESULT 15
CNS0060N/C 910 bp DNA linear GSS 03-JUN-1999
LOCUS BACR14J21 of RPCT-98 library from Drosophila melanogaster (fruit fly) genomic survey sequence.
DEFINITION BACR14J21 of RPCT-98 library from Drosophila melanogaster (fruit fly) genomic survey sequence.
ACCESSION AL065629.1 GI:4944698
VERSION AL065629
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 910)
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segefe@genoscope.cns.fr)
Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oosawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCT-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1..910
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14J21"
/clone_lib="RPCT-98"
/note="end : T7"

ORIGIN

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 14:33:42 ; Search time 5602.61 seconds

(without alignments)
11294.891 Million cell updates/sec

Title: US-10-076-747-8

Perfect score: 1460

Sequence: 1 ggcctggcctcgtctatccc.....ggcagcagagccacgcagct 1460

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 2167151995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_da:*
2: gb_da:*
3: gb_da:*
4: gb_da:*
5: gb_da:*
6: gb_da:*
7: gb_da:*
8: gb_da:*
9: gb_da:*
10: gb_da:*
11: gb_da:*
12: gb_da:*
13: gb_da:*
14: gb_da:*
15: gb_da:*
16: gb_da:*
17: gb_da:*
18: gb_da:*
19: gb_da:*
20: gb_da:*
21: gb_da:*
22: gb_da:*
23: gb_da:*
24: gb_da:*
25: gb_da:*
26: gb_da:*
27: gb_da:*
28: gb_da:*
29: gb_da:*
30: gb_da:*
31: gb_da:*
32: gb_da:*
33: gb_da:*
34: gb_da:*
35: gb_da:*
36: gb_da:*
37: gb_da:*
38: gb_da:*
39: gb_da:*
40: gb_da:*
41: gb_da:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	642.2	44.0	755	9	HSIGVL033
2	634.2	43.4	895	9	BC033102
3	626	42.9	922	9	BC062711
4	607	41.6	891	6	AK135365
5	606.8	41.6	811	9	AB064153
6	600.4	41.1	789	9	AB064207
7	597.4	40.9	877	9	HUMIGHPEAL
8	591.4	40.5	1230	9	AK130519
9	588.4	40.3	894	9	BC018749
10	584.6	40.0	756	9	HSIGVL034
11	584.4	40.0	810	9	AB064158
12	581.4	39.8	915	9	BC015833
13	570.2	39.1	783	9	HSIGVL001
14	570	39.0	806	9	AB064155
15	570	39.0	810	9	AB064163
16	568.4	38.9	807	9	AB064154
17	568.4	38.9	808	9	AB064164
18	568.4	38.9	916	9	BC030983
19	568.2	38.9	812	9	AB064160
20	566.8	38.8	808	9	AB064152
21	565.4	38.7	729	9	AY172960
22	563.6	38.5	809	9	AB064156
23	562	38.5	805	9	AB064161
24	560.2	38.4	810	9	AB064159
25	559	38.3	885	9	AK129585
26	557	38.2	826	9	AB064215
27	550.6	37.7	807	9	AB064162
28	544.8	37.3	808	9	AB064157
29	543	37.2	887	9	AK130461
30	537.6	36.8	663	9	HS007991
31	531.6	36.4	885	9	AK129603
32	520.4	35.6	915	9	BC012876
33	518.4	35.5	942	6	AX780304
34	511.8	35.1	2112	6	AX523520
35	510.6	35.0	666	9	HS007992
36	500.2	34.3	903	9	BC020233
37	499.8	34.2	919	9	BC022098
38	498.4	34.1	935	6	AR135362
39	498.2	34.1	819	9	HSIGVL029
40	497.2	34.1	605	6	AX379222
41	489.6	33.5	9472	6	AX287808
42	485.4	33.2	652	9	HUMIGHVXC
43	485.2	33.2	816	9	AB064145
44	482	33.0	870	6	BD176841
45	482	33.0	870	9	HSIGLV

ALIGNMENTS

RESULT 1
LOCUS HSIGVL033
DEFINITION Human rearranged immunoglobulin lambda light chain mRNA.
ACCESSION X57823
VERSION X57823.1 GI.33745
KEYWORDS Ig lambda light chain; immunoglobulin.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 344 to 390)
V Combarato, G. and Klobbeck, H.G.
TITLE
Immunoglobulin lambda light chain gene segments of the human
immunoglobulin lambda light chain locus are separated by 14 kb and

rearrange by a deletion mechanism
 JOURNAL Eur J Immunol. 21 (6), 1513-1522 (1991)
 MEDLINE 91257162
 PUBMED 1904362
 REFERENCE 2 (bases 1 to 755)
 AUTHORS Klobbeck, H.G.
 TITLE Submitted
 JOURNAL Submitted (31-JAN-1991) H.G. Klobbeck, Inst fuer Physiologische
 Chemie, Physiologische Biochemie und Zellbiologie der Universitaet
 Muenchen, Shillerstr. 44, 8000 Muenchen 2, Germany
 COMMENT for overlapping sequences see: X51754-55; J00252-54; M15641-42.
 FEATURES
 source
 1..755
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /isolate="individual S"
 /db_xref="taxon:9606"
 /chromosome="22"
 /clone="CS105"
 /cell_type="B-cell"
 /tissue_type="spleen"
 /clone_id="phage library CS"
 1..755
 /gene="immunoglobulin lambda light chain"
 1..755
 /gene="immunoglobulin lambda light chain"
 /note="cDNA"
 /evidence="experimental"
 1..708
 /gene="immunoglobulin lambda light chain"
 /protein_id="CAA0960.1"
 /db_xref="GI:33746"
 /db_xref="REMBL:CAA0960"
 /translation="MAVALFRLTLTGSGTSMASALTPASVSGSPGSSITISGCS
 SDGCVNVVSNKQOHPKAKIMTYDTPRPSVSRPFGSGSAGTASITISGDE
 DEADYCTSKTSSTSYVGTGKYSVLDGPANFTVLFPSSSELDANRATVCLIS
 DFGAVTAVAKADGSPVAKGVETTKSKSNKNRYAASYSILSTPEQMKSHRSYCV
 TKGSTVEKTVAFTECS"
 1..57
 /gene="immunoglobulin lambda light chain"
 58..388
 /gene="immunoglobulin lambda light chain"
 /note="variable region
 V(lambda)II"
 351..388
 /gene="immunoglobulin lambda light chain"
 /note="J-segment"
 389..708
 /gene="immunoglobulin lambda light chain"
 /note="constant region"
 ORIGIN
 Query Match 44.0%; Score 642.2; DB 9; Length 755;
 Best Local Similarity 96.5%; Pred. No. 7e-136;
 Matches 667; Conservative 0; Mismatches 23; Indels 1; Gaps 1;
 Db 1
 QY 1 GGCCTGGGCTCTGCTATTCCTCACTCTCTCACTAGGAGCAGAGGCTCTGGGCCCACTC 60
 Db 3 GGCCTGGGCTCTGCTATTCCTCACTCTCTCACTAGGAGCAGAGGCTCTGGGCCCACTC 62
 QY 61 TGCCTGACTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 Db 63 TGCCTGACTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 122
 QY 121 CACTGGAACAGAGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 Db 123 CACTGGAACAGAGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 182
 QY 181 AGGCAAGGCCCCCAACTCATATTATGAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 240
 Db 183 AGGCAAGGCCCCCAACTCATATTATGAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 242

QY 241 TGCCTTCTGCTGCTCCAACTCTGSCAACAGGAGCTCTCTGACACATCTCTGGGCTCCAGGC 300
 Db 243 TGCCTTCTGCTGCTCCAACTCTGSCAACAGGAGCTCTCTGACACATCTCTGGGCTCCAGGC 302
 QY 301 TGAAGAGAGAGCTGATTATTACTGCTGCTCAATATCAAGAACTACTTCTCATGCTTCCG 360
 Db 303 TGAAGAGAGAGCTGATTATTATTGACACCTCAAAAACAGACACACTTTATGCTTCCG 362
 QY 361 AACTGGACCAAGGTCACGCTCTTGTAGTACGCCCAAGCCCAACCCCATGCTACTCTGT 420
 Db 363 GACTGGACCAAGGTCCTCGTCTGTAGTACGCCCAAGCCCAACCCCATGCTACTCTGT 422
 QY 421 CCGGCTCTCTGAGAGAGCTCCAAAGCCCAAGGCTCACTAGTGTCTGATCAGTGA 480
 Db 423 CCGGCTCTCTGAGAGAGCTCCAAAGCCCAAGGCTCACTAGTGTCTGATCAGTGA 482
 QY 481 CTCTTACCCGGAGAGCTGTGACAGTGGCTGGAAGGCAATGGACCCCGTCAAGCGGG 540
 Db 483 CTCTTACCCGGAGAGCTGTGACAGTGGCTGGAAGGCAATGGACCCCGTCAAGCGGG 542
 QY 541 AGTGGAGACACCAAAACCTCCAAACAGAGCAAAACAAAGTACGAGGCGACAGCTACT 600
 Db 543 AGTGGAGACACCAAAACCTCCAAACAGAGCAAAACAAAGTACGAGGCGACAGCTACT 602
 QY 601 GAGCTGAGCGCCGAGAGTGAAGTCCCAAGAGCTACAGCTCCAGTCCAGTCACGATGA 660
 Db 603 GAGCTGAGCGCCGAGAGTGAAGTCCCAAGAGCTACAGCTCCAGTCCAGTCACGATGA 662
 QY 661 AGGAGACACCTGTGATGAAGACAGTGTGCC 691
 Db 663 AGGAGACACCTGTGATGAAGACAGTGTGCC 692
 RESULT 2
 BC033102 895 bp mRNA linear PRI 19-NOV-2003
 LOCUS Homo sapiens cDNA clone MGC:45681 IMAGE:4851128, complete cds.
 DEFINITION BC033102
 ACCESSION BC033102.1 GI:21619847
 VERSION
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 895)
 REFERENCE
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenfer, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
 Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Umedin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulvik, S.W.,
 Villalón, D.K., Wuzny, D.M., Sodergren, E., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, D., Kettman, M., Madan, A., Rodriquez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Krzyzanski, M.I., Skalka, U., Smalins, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 MEDLINE 22388257
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 895)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (25-JUN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer

REMARK

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Louis Staudt
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdel, Yaron Buterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywnski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Matheson, Candice McLeavy, Steven Nees, Pawan Pandoh, Anna-Lies Prabhu, Parvaneh Saeedi, Jacqueline Schein, Diane Smalhus, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Mizanada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.lnl.gov>
 Series: TRL Plate: 43 Row: 9 Column: 22
 This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction, similarity but not identity to protein.

FEATURES

source

Location/Qualifiers
 1..895
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:45681 IMAGE:4851128"
 /tissue_type="Primary B-Cells from Tonsils"
 /clone_id="NIH MGC_48"
 /lab_host="MDH105-R"
 /note="vector: pOTB7"
 31..738
 /codon_start=1
 /product="unknown (protein for MGC:45681)"
 /protein_id="AAH3102.1"
 /db_xref="GI:21619848"
 /translation="MAKALLLTLLTGSGMAQALTOPPSAGSGSGVTSCTGT
 SSDVGYNYSHYQHPKAKPKLTYVNNRPSGVPRSGSSGNTATVGLAE
 DEADYCSYAGSNVYFGTGTATVGLGPKAPFTVTLPPSSSELDQANATVCLIS
 DFYGAIVAMKADGSPVKAQVETTKPSKSNKRYASVLSLTPQWMSHRSYSCOV
 THESGTEKTVAPTECS"

CDS

misc_feature

/note="IGV; Region: Immunoglobulin domain variable region
 (v) subfamily"
 /db_xref="CDD:c00099"
 436..717
 /note="IGC; Region: Immunoglobulin domain constant region
 subfamily"
 /db_xref="CDD:c00098"

misc_feature

Query Match 43.4%; Score 634.2; DB 9; Length 895;
 Best Local Similarity 95.8%; Pred. No. 4,7e-134;
 Matches 662; Conservative 0; Mismatches 28; Indels 1; Gaps 1;

ORIGIN
 1 GAGCGGGGCTGCTATCTCCACCTCTCACTCAGGAGGAGGCTGGGCGCCAGTC 60
 33 GAGCGGGGCTGCTATCTCCACCTCTCACTCAGGAGGAGGCTGGGCGCCAGTC 92
 61 TGCCCTGACTGAGTGTGCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 93 TGCCCTGACTGAGTGTGCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 152
 121 CACTGGAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 180
 153 CACTGGAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 212
 181 AGGCAAAAGCCCAAACTCATTTATGAGTGACGATGATGCGCCCTCAGGGGTTCTTA 240

DB 213 AGGCAAAAGCCCAAACTCATTTATGAGTGACGATGATGCGCCCTCAGGGGTTCTTGA 272
 QY 241 TCGCTTTCTGTGCTCCAAAGTGTGGCAACAAGGCTCTCTGACCATCTCTGGCTCCAGGC 300
 DB 273 TCGCTTTCTGTGCTCCAAAGTGTGGCAACAAGGCTCTCTGACCATCTCTGGCTCCAGGC 332
 QY 301 TGAGGACGAGGCTGATTTATTTATGCTGCTCATATCAAGAGTATTTGATGCTTCCG 360
 DB 333 TGAGGATGAGGCTGATTTATTTATTTATGCTGCTCATATCAAGAGTATTTGATGCTTCCG 392
 QY 361 AACTGGACCAAGTCAACGCTGCTGATGCTGACGCCCAAGCCCAAGCCCTGCTGCTGTT 420
 DB 393 AACTGGACCAAGTCAACGCTGCTGATGCTGACGCCCAAGCCCAAGCCCTGCTGCTGTT 452
 QY 421 CCGCGCTCTCTTGTAGAGCTCCAAAGCCCAAGCCCAAGCCCTGCTGCTGATGATGTA 480
 DB 453 CCGCGCTCTCTTGTAGAGCTCCAAAGCCCAAGCCCAAGCCCTGCTGCTGATGATGTA 512
 QY 481 CTTCTAACCGGAGGCTGTGACAGTGTGCTGGAAGGCAATGCGACCCCTGCAAGCGGG 540
 DB 513 CTTCTAACCGGAGGCTGTGACAGTGTGCTGGAAGGCAATGCGACCCCTGCAAGCGGG 572
 QY 541 AGTGAAGACCAACCAACCTCTCAACCAAGCAACCAAGTACGCGGCCAGCAGCTTACT 600
 DB 573 AGTGAAGACCAACCAACCTCTCAACCAAGCAACCAAGTACGCGGCCAGCAGCTTACT 632
 QY 601 GAGCCTGACGCCCGGACGATGGAAGTCCCAAGAGTCAAGCTGCTGCTGACGACATGA 660
 DB 633 GAGCCTGACGCCCGGACGATGGAAGTCCCAAGAGTCAAGCTGCTGCTGACGACATGA 692
 QY 661 AGGAGCAACCGTGATGATGAAGACAGTGTGCTCC 691
 DB 693 AGGAGCAACCGTGATGATGAAGACAGTGTGCTCC 722

RESULT 3

LOCUS BC062711 922 bp mRNA linear PRI 26-NOV-2003
 DEFINITION Homo sapiens cDNA clone MGC:72002 IMAGE:30350364, complete cds.
 ACCESSION BC062711
 VERSION BC062711.1 GI:38541896
 KEYWORDS MGC.

SOURCE

ORGANISM

Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 922)

REFERENCE

AUTHORS

Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner J., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhac N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udell T.B., Toschiviaki S., Carrinci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Wulianhy S.V., Bosak S.A., McSwain P.U., McKernan R.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliyil S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Scheraga A., Schein J.E., Jones S.J. and Marra M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 922)
 Strusberg R.

TITLE JOURNAL

Direct Submission
Submitted (25-NOV-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Heiao, Martin Krzyminski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Seedei, Jacqueline
Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Scott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAL Place: 51 Row: b Column: 14
This clone was selected for full length sequencing because it
passed the following selection criteria: Genomescan gene
prediction, similarity but not identity to protein.

FEATURES source

1..922
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:72002 IMAGE:30350364"
/tissue_type="Glandular pool- thyroid, parathyroid,
adrenal cortex, pineal gland, submandibular gland."
/clone_11b="Nih MGC_184"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
47..757
/codon_start=1
/product="Unknown (protein for MGC:72002)"
/protein_id="AAH6711.1"
/db_xref="GI:38541897"
/translation="MAVALLLILLTQTGSGMAQSALTQPSVSGSPQSITISCTG
SSDVGYDFVSMYQONPGKAPKLIIVDSNPSGVSFSGSKGNATSLITGLGLE
DEADYVGSYSSTTPFVFGSGTKVTVGGKAPVTYLPPESEELQANKATVLCI
SDRYGATVAMKADGSPVKAGVETTKSKSNKKVAASVLSLTPEQWMSHSYSCQ
VTHSGTYEKTVAPEKCS"

misc_feature

131..436
/note="IGV; Region: Immunoglobulin domain variable region
(v) subfamily"
/db_xref="CDI:cd00099"
455..736
/note="IGC; Region: Immunoglobulin domain constant region
subfamily"
/db_xref="CDI:cd00098"

ORIGIN

Query Match 42.9%; Score 626; DB 9; Length 922;
Best Local Similarity 95.8%; Pred. No. 3,5e-133;
Matches 665; Conservative 0; Mismatches 25; Indels 4; Gaps 2;
QY 1 GGCCTGGGCTGTGTAATCTCAACCTCTCTCACTAGGAGCAAGAGGTCTGGGCCCAATC 60
DB 49 GGCCTGGGCTGTGTAATCTCAACCTCTCTCACTAGGAGCAAGAGGTCTGGGCCCAATC 108
QY 61 TGCCCTGACTAGTGTGCTCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 120
DB 109 TGCCCTGACTAGTGTGCTCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 168
QY 121 CACTGAAACGACGACTACGTTGTGTATTAATACTATGTCTCTGTGATCAACAACGACCC 180

DB 169 CACTGAAACGACGACTACGTTGTGTATTAATACTATGTCTCTGTGATCAACAACCC 228
QY 181 AGGCAAGCCCCCAACTCATATTATATAGGTAGTAATCGGCCCTCAGGGGTTCTTA 240
DB 229 AGGCAAGCCCCCAACTCATATTATATAGGTAGTAATCGGCCCTCAGGGGTTCTCA 288
QY 241 TGCCCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 300
DB 289 TGCCCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 348
QY 301 TGAAGCAGAGGCTGATTTACTGTCTGTCTCATATACAAAGTAAGTACT--TCTCATGTCTT 357
DB 349 TGAAGCAGAGGCTGATTTACTGTCTGTCTCATATACAAAGTACT--TCTCATGTCTT 408
QY 358 CGGAATGTGAGACCAAGGTACCGTCTGTAGTCAAGCCCAAGCCCAACCCCATGTCACTT 417
DB 409 CGGATCTGGGACCAAGGTACCGTCTGTAGTCAAGCCCAAGCCCAACCCCATGTCACTT 468
QY 418 GTTCCCGCCCTCTGTGAGAGCTCCAGGCAAGGCAACAGTGTGTCTATGAG 477
DB 469 GTTCCCGCCCTCTGTGAGAGCTCCAGGCAAGGCAACAGTGTGTCTATGAG 528
QY 478 TGAATTTCAACCCGAGAGCTGTGACAGTGGCTGGAAGGCAATGGCAAGCCCGTCAAGGC 537
DB 529 TGAATTTCAACCCGAGAGCTGTGACAGTGGCTGGAAGGCAATGGCAAGCCCGTCAAGGC 588
QY 538 GGAAGTGGAGACCAACCAACCTCCCAACAGAGCAACCAAGTACGGGCGAGAGACTA 597
DB 589 GGAAGTGGAGACCAACCAACCTCCCAACAGAGCAACCAAGTACGGGCGAGAGACTA 648
QY 598 CTTGACCTGAGCCCGAGAGCTGTGAGAGTCCCAAGAACTACAGCTGCAGGTACGCA 657
DB 649 CTTGACCTGAGCCCGAGAGCTGTGAGAGTCCCAAGAACTACAGCTGCAGGTACGCA 708
QY 658 TGAAGGAGACACCTGTGATGAGAGACAGTGTCTCC 691
DB 709 TGAAGGAGACACCTGTGATGAGAGACAGTGTCTCC 741

RESULT 4
AR135365
LOCUS AR135365 891 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 23 from patent US 6135941.
ACCESSION AR135365
VERSION AR135365.1 GI:14476037
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 891)
Hillman,J.L., Lal,P., Tang,Y.Tom., Yue,H., Au-Young,J.,
Corley,N.C., Guegler,K.J. and Baughn,M.R.
Human immune system associated molecules
Patent: US 6135941-A 23 24-OCT-2000;
TITILE Location/Qualifiers
JOURNAL 1..891
FEATURES
source /organism="Unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 41.6%; Score 607; DB 6; Length 891;
Best Local Similarity 93.3%; Pred. No. 7.6e-128;
Matches 645; Conservative 0; Mismatches 45; Indels 1; Gaps 1;
QY 1 GGCCTGGGCTGTGTAATCTCAACCTCTCTCACTAGGAGCAAGAGGTCTGGGCCCAATC 60
DB 36 GGCCTGGGCTGTGTAATCTCAACCTCTCTCACTAGGAGCAAGAGGTCTGGGCCCAATC 95
QY 61 TGCCCTGACTAGTGTGCTCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 120
DB 96 TGCCCTGACTAGTGTGCTCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 155

QY 121 CACTGGAACAGACGACGCTGGTGGTAACTATGCTCTGTAACAGACAGACC 180
 DB 156 CACTGGAACAGACGACGCTGGTGGTAACTATGCTCTGTAACAGACAGACC 215
 QY 181 AGGCAAAAGCCCAACTCATATTTATGAGTCAAGTATCGGCGCTCAGGGTTTCTAA 240
 DB 216 AGGCAAGGCGCCCAACTCATATTTATGAGTCAAGTATCGGCGCTCAGGGTTTCTAA 275
 QY 241 TCGCTTCTGCTGCTCCAGTCTGAGCAACAGGCGCTCCCTCAACATCTTGGGCTCCAGGC 300
 DB 276 TCGGTTCTGCTGCTCCAGTCTGAGCAACAGGCGCTCCCTCAACATCTTGGGCTCCAGGC 335
 QY 301 TGAAGACAGAGCTGATTATTACTGCTCTCAATATACAGAGTACTTCTCATGCTTCCG 360
 DB 336 TGAAGACAGAGCTGATTATTACTGCTCTCAATATGAGCAACAGTCTGCTTCCG 395
 QY 361 AACTGGGACCAAGTCAACCGTCTGAGTCAAGCCCAAGCCCACTGTCACCTGTT 420
 DB 396 CGAGGAGCAAGTCAACCGTCTGAGTCAAGCCCAAGCCCACTGTCACCTGTT 455
 QY 421 CCGGCGCTCTCTGAGGAGCTCCAGGCAACAGGCGCACTAGTGTCTGATCAGTGA 480
 DB 456 CCGGCGCTCTCTGAGGAGCTCCAGGCAACAGGCGCACTAGTGTCTGATCAGTGA 515
 QY 481 CTCTTACCCGAGCTGTGACAGTGGCTGGAAGGAGGAGGAGGCGGCTCAAGGCGG 540
 DB 516 CTCTTACCCGAGCTGTGACAGTGGCTGGAAGGAGGAGGAGGCGGCTCAAGGCGG 575
 QY 541 AGTGAAGACCAACAACTCTTCAACAGAGCAACAGTATGCGGCGGAGAGTAACT 600
 DB 576 AGTGAAGACCAACAACTCTTCAACAGAGCAACAGTATGCGGCGGAGAGTAACT 635
 QY 601 GAGGCTGACGCGCCAGGAGTGAAGTCCACAGAGTATGAGGAGGAGGAGGAGTGA 660
 DB 636 GAGGCTGACGCGCCAGGAGTGAAGTCCACAGAGTATGAGGAGGAGGAGGAGTGA 695
 QY 661 AGGAGCAACCGTGTATGAAGACAGTGTCC 691
 DB 696 AGGAGCAACCGTGTGA-GAAGCAGTGTCC 725
 RESULT 5
 AB064153
 LOCUS Homo sapiens IGL mRNA for immunoglobulin lambda light chain VLJ
 DEFINITION
 AB064153
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 1 Akhori, Y., Iba, Y., Morino, K., Shinohara, M., Hirono, Y., Kakita, M.,
 Suzuki, K., Torii, H., Ueki, Y., Honda, T., Katsunuma, H., Okada, J.,
 Miura, K. and Kurosawa, Y.
 Construction and characterization of antibody libraries: isolation
 of therapeutic human antibodies and application to functional
 genomics
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Kurosawa, Y.
 Direct Submission
 Submitted (25-JUN-2001) Yoshikazu Kurosawa, Institute for
 Comprehensive Medical Science, Fujita Health University,
 Kutsurake-cho, Toyosake 470-1192, Japan
 (E-mail: kurosawa@fujita-hu.ac.jp, Tel: 81-562-93-9387)
 COMMENT
 Please visit our web site
 URL: http://www.fujita-hu.ac.jp/immunity/
 Location/Qualifiers
 1..811
 /organism="Homo sapiens"

/mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="AIMS4"
 /note="Mixture of tissues: tonsils, umbilical cords,
 peripheral blood and bone marrow"
 1..811
 /gene="IGL"
 1..811
 /gene="IGL"
 /codon_start=1
 /product="immunoglobulin lambda light chain VLJ region"
 /protein_id="BAC01781.1"
 /db_xref="GI:21669513"
 /translation="MKYLLPTAAAGLLIAAPAMAGSAITOPASVSGPSGQITISG
 TGISDVGIVIVSWYQHPKAKLMIYDKSPSGVSNFPGSKSNTSLTISGL
 QAEEDADYCSYSSSTYYVGTGRTKYLQAPANPVTLFPSSSELQNKATLVC
 LISDFYPAIVAMKADSPYKAGVETKPSKQNNKXVASSYSLTPEQKSHRSYS
 CVTHGSEVETKVAPTFESARQSTPFVCEYQSSDLPQPPNVAAGSGGG"
 1..66
 /gene="IGL"
 /note="pep1b signal peptide"
 sig_peptide
 ORIGIN
 Query Match 41.6%; Score 606.8; DB 9; Length 811;
 Best Local Similarity 97.2%; Pred. No. 8.4e-128;
 Matches 628; Conservative 0; Mismatches 17; Indels 1; Gaps 1;
 QY 46 GTCTGGGCGCAAGTCTGCTGCTGACTCACTGCTGCTGCTGCTGCTGCTGCTG 105
 DB 57 GGCCTAGGCGCAAGTCTGCTGCTGACTCACTGCTGCTGCTGCTGCTGCTGCTG 116
 QY 106 GATCACCATCTCTGCACTGGAACAGAGTCAAGTGTGTTATATCTATGTCCTG 165
 DB 117 GATCACCATCTCTGCACTGGAACAGAGTCAAGTGTGTTATATCTATGTCCTG 176
 QY 166 GTACCAACAGACCCAGGCAAAAGCCCAACTATCTATTATGAGTCAAGTATG 225
 DB 177 GTACCAACAGACCCAGGCAAAAGCCCAACTATCTATTATGAGTCAAGTATG 236
 QY 226 CTCAGGGGTTTCTATAGCTCTCTGAGTCTGCAAGTCTGCAAGCGGCTCCGACAT 285
 DB 227 CTCAGGGGTTTCTATAGCTCTCTGAGTCTGCAAGTCTGCAAGCGGCTCCGACAT 296
 QY 286 CTCGAGGCTCCAGGCTGAGAGCAAGGCTGATTATTAAGTCTGCTCATATCAAGAGTAC 345
 DB 297 CTCGAGGCTCCAGGCTGAGAGCAAGGCTGATTATTAAGTCTGCTCATATCAAGAGTAC 356
 QY 346 TTCTCATGTTCTTGGAACTGGGACCAAGTCAAGCTCTAGTCAAGCCCAAGCCCAAGCC 405
 DB 357 CACTTATGTTCTTGGAACTGGGACCAAGTCAAGCTCTAGTCAAGCCCAAGCCCAAGCC 416
 QY 406 CACTGCTCACTGTGTTCCGCGCTCTCTGAGAGCTCCCAAGCAAGGCGCACTAGT 465
 DB 417 CACTGCTCACTGTGTTCCGCGCTCTCTGAGAGCTCCCAAGCAAGGCGCACTAGT 476
 QY 466 GTGTGTATCAAGTCACTTCAACCGGAGCTGTGACAGTGGCTTGAAGAGGAGTGGCAG 525
 DB 477 GTGTGTATCAAGTCACTTCAACCGGAGCTGTGACAGTGGCTTGAAGAGGAGTGGCAG 536
 QY 526 CCGGCTCAAGGCGGAGAGTGAACCAACCAACCGTCCCAACAGAGCAACCAAGTACGC 585
 DB 537 CCGGCTCAAGGCGGAGAGTGAACCAACCAACCGTCCCAACAGAGCAACCAAGTACGC 596
 QY 586 GGCAGAGCTTACTGAGCTGAGCGCCCGAGAGTGAAGTCCCAAGAGCTTACAGCTG 645
 DB 597 GGCAGAGCTTACTGAGCTGAGCGCCCGAGAGTGAAGTCCCAAGAGCTTACAGCTG 656
 QY 646 CCAAGTCAAGATGAAGGAGACCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 691
 DB 657 CCAAGTCAAGATGAAGGAGACCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 701

RESULT 6
AB064207 789 bp mRNA linear PRI 02-JUL-2002
LOCUS Homo sapiens IGL mRNA for immunoglobulin lambda light chain VLJ
DEFINITION
AB064207 region, partial cds, clone:167.
ACCESSION
AB064207.1 GI:21669620
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS Akahori, Y., Iba, Y., Morino, K., Shinohara, M., Hiroo, Y., Kakita, M.,
1 Suzuki, K., Torii, H., Ueki, Y., Honda, T., Katsun, H., Okada, Y.,
Miyata, K. and Kurosawa, Y.
TITLE Construction and characterization of antibody libraries: isolation
of therapeutic human antibodies and application to functional
genomics
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 789)
AUTHORS Kurosawa, Y.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-2001) Yoshikazu Kurosawa, Institute for
Comprehensive Medical Science, Fujita Health University;
Kutsukake-cho, Toyosake 470-1192, Japan
COMMENT Please visit our web site
URL: http://www.fujita-hu.ac.jp/immunity/
FEATURES
source
1..789
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="167"
/clone_1ib="ALMS4"
/note="mixture of tissues: tonsils, umbilical cords,
peripheral blood and bone marrow"
1..789
/gene="IGL"
/gene="IGL"
1..789
/gene="IGL"
/codon_start=1
/product="immunoglobulin lambda light chain VLJ region"
/protein_id="BAC01835.1"
/db_xref="GI:21669621"
/translation="MKYLLPTAAAGLILIAQPAAGSVLTLPASVSGSGPGQSITISG
NQTSDVGVYVSWYQOHGKAPKLMYDYKSPSGVSNRPSGSKGNPASTISGL
QAEEDADYICSSYSSSTVFTGTGKTVLQPKAPVTVTFPSSSELDANKATLYC
LISDFPAVTVAMADSPVAGVETTPKSGSNKXKAAASYLTPEDQKSHRSYS
CVTHEGSTVEKTVAPTECSARQSTPFVCEYQOGSDLPDPVNA"
1..66
/gene="IGL"
/note="pelb signal peptide"

Db 237 CTCAGGGGTTCTTAATCGCTTCTGTGCTCCAAAGCTGGCAACAGGCTCCCTGACCAT 296
Qy 286 CTCAGGGCTCCAGGCTGAGAGAGAGAGCTGATATTACTGCTCTCATATACAGAGTAC 345
Db 297 CTCAGGGCTCCAGGCTGAGAGAGAGAGCTGATATTACTGCTCTCATATACAGAGTAC 356
Qy 346 TTCTCATGTCCTTGGAACTGGGACCAAGGTCAACCGTCTAGTCAAGCCCAAGCCCAACC 405
Db 357 CACTATGTCCTTGGAACTGGGACCAAGGTCAACCGTCTAGTCAAGCCCAAGCCCAACC 416
Qy 406 CACTGTCATGTCCTTGGAACTGGGACCAAGGTCAACCGTCTAGTCAAGCCCAAGCCCAACC 465
Db 417 CACTGTCATGTCCTTGGAACTGGGACCAAGGTCAACCGTCTAGTCAAGCCCAAGCCCAACC 476
Qy 466 GTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 525
Db 477 GTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 536
Qy 526 CCCCCCAAGGGGGAGTGGAGACCAACCAACCCCTCAACAGAGCAACAGTACAGCC 585
Db 537 CCCCCCAAGGGGGAGTGGAGACCAACCAACCCCTCAACAGAGCAACAGTACAGCC 596
Qy 586 GGCACAGCTACCTGAGCTGAGCCGCGAGAGTGAAGTCCACAGAGCTACAGCTG 645
Db 597 GGCACAGCTACCTGAGCTGAGCCGCGAGAGTGAAGTCCACAGAGCTACAGCTG 656
Qy 646 CACAGTCACGATGAAGAGAGACCGGTGATGAAGAGTGTCC 691
Db 657 CACAGTCACGATGAAGAGAGACCGGTGATGAAGAGTGTCC 701

RESULT 7
HUMICHEPAL 877 bp mRNA linear PRI 08-OCT-1993
LOCUS Human (hybridoma H210) anti-hepatitis A immunoglobulin lambda chain
DEFINITION
variable region, constant region, complementarity-determining
regions mRNA, complete cds.
ACCESSION
VERSION M87790.1 GI:185363
KEYWORDS complementarity-determining region; constant region;
immunoglobulin lambda chain; light chain; variable region.
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Lewis, A.P., Parry, N., Peakman, T.C. and Crowe, J.S.
JOURNAL Unpublished (1992).
REFERENCE
AUTHORS Lewis, A.P., Lemon, S.M., Barber, K.A., Murphy, P., Parry, N.R.,
Peakman, T.C., Sime, M.J., Worden, J. and Crowe, J.S.
TITLE Rescue, expression, and analysis of a neutralizing human
anti-hepatitis A virus monoclonal antibody
JOURNAL J. Immunol. 151 (5), 2829-2838 (1993)
MEDLINE 93367243
PUBMED 8395549
COMMENT Original
FEATURES
source
5' UTR
CDS
1..877
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/note="anti-hepatitis A; putative"
1..739
/note="putative"
32..739
/note="anti-hepatitis A; putative"
/product="immunoglobulin lambda chain"
/protein_id="AA02915.1"
/db_xref="GI:185364"
/translation="MAMALLLTLLTODTGSMAQALTPASVSGSGPGQSITISGCTGT
NNDVDSYNLVSWYQOHGKAPKLMYDYKSPSGVSNRPSGSKGNPASTISGLQAE

Query Match	Similarity	Score	DB	Length
Best Local	92.5%	597.4	9	877
Matches	639	Conservative	0	Mismatches 51; Indels 1; Gaps 1;

Query	Match	Similarity	Score	DB	Length
QY	1	GGCCTGGAGCTCTGCTATTCTTCACCCCTCTCACTCAGGGCACAGGGTCTTGAGCCCACTG	60		
Db	34	GGCTCTGGAGCTCTGCTGCTCTTCACCCCTCTCACTCAGGACAGAGGTCCTTGAGCCCACTG	93		
QY	61	TGCCCTGACTCAGTCTGCTGCTCGATGTCGGGTCTCCTGAAACAGTCATCACTCTCTG	120		
Db	94	TGCCCTGACTCAGCTGCTCGATGTCGGGTCTCCTGAAACAGTCATCACTCTCTG	153		
QY	121	CACTCGAAACCAAGCACTACGTTTGGGTATTAACTATGTCCTCTGATCCAAACAGCACCC	180		
Db	154	CACTCGAAACCAACATGATGTTGGAGTTATPAACCTTATCTCTGATCCAAACAGCACCC	213		
QY	181	AGGCAAAAGCCCAAACTCATCTATTATGAGGTCAGTAATCGGCTCTCAGGGTTTCTAA	240		
Db	214	AGGCAAAAGCCCAAAATCATGATTTATGAGGTCAGTAAGCGGCTCTCAGGGTTTCTAA	273		
QY	241	TGCGTCTCTGTGGCTCCAGTCTGGAACACAGGCGCCCTGACATCTCTGAGGCTCCAGGC	300		
Db	274	TGCGTCTCTGTGGCTCCAGTCTGGAACACAGGCGCTCTGACATCTCTGAGGCTCCAGGC	333		
QY	301	TGAGGACAGGCTGATTATTACTGCTGCTCATATACAAAGATACCTTCATGTCCTCCG	360		
Db	334	TGAGGACAGGCTGATTATTACTGCTGCTCATATGCAAGTATGATACATGATGATTTCCG	393		
QY	361	AACGTGGAGCAAGGCAACGATCTTAAGTCCAGGCCCAAGGCCCACTGTCATCTCTTT	420		
Db	394	CGAGAGGACCAAACTGACCGTCTTAAGTCCAGGCCCAAGGCCCACTGTCATCTCTTT	453		
QY	421	CCCGCGCTCTCTGAGAGCTTCCAAAGCCAAAGGCCCACTAGTGTCTTGATCATGTA	480		
Db	454	CCCGCGCTCTCTGAGAGCTTCCAAAGCCAAAGGCCCACTAGTGTCTTCATPAAGTA	513		
QY	481	CTTCTAACCGGAGGTGTGACAGTGGCTGGAAGGCATATGAGACCCCTCTCAAGCGGG	540		
Db	514	CTTCTAACCGGAGGTGTGACAGTGGCTTGAAGGCATATGAGACCCCTCTCAAGCGGG	573		
QY	541	AGTGAAGACCAACCAACCTCTCCAAACAGAGCAACAACAAGTACGAGGCAACAGTAACT	600		
Db	574	AGTGAAGACCAACCAACCTCTCCAAACAGAGCAACAACAAGTACGAGGCAACAGTAACT	633		

RESULT 8	
AK130519	
LOCUS	AK130519
DEFINITION	Homo sapiens CDNA FLJ27009 f.1s, clone SLV051c1, highly similar to
ACCESSION	AK130519
VERSION	AK130519.1 GI:34527337
KEYWORDS	oligo capping; f.1s (full insert sequence).
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

Nishikawa, T., Sugiyama, A., Kawakami, B., Negai, K., Isogai, T. and Sugano, S.

JOURNAL REFERENCE AUTHORS TITLE	Unpublished
2 (baees 1 to 1230)	
Sugano, S. and Suzuki, Y.	
Direct Submission	
Submitted (31-JUN-2003)	Sumio Sugano, Institute of Medical Science,

COMMENT NEDO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert sequencing
 Research Association for Biotechnology (RAB); cDNA library
 construction and 5'-end one pass sequencing: Institute of Medical
 Science, University of Tokyo; Laboratory of genome Structure, Human
 Genome Center; 3'-end one pass sequencing: RAB; clone selection for
 full insert sequencing: RAB and Helix Research Institute.

FEATURES	location/Qualifiers
source	1. .1230

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SLV05161"
/tissue_type="salivary gland"
/clone_id="SLV"
/note="cloning vector: pME18SFL3
42..755

```

[illegible]

	ORIGIN				
	Query Match	40.5%	Score 591.4;	DB 9;	Length 1230;
	Best Local Similarity	89.1%;	Pred. No. 2,7e-124;		
	Matches 687; Conservative	0; Mismatches 71;	Indels 13; Gaps		4
QY	1 GGCGCTGGAGCTGTGATTCCTCAACCTCCTCAACTCAGGAGCACAGGAGTCTCTGGCCCAAGTC	60			

Query Match 40.3%; Score 588.4; DB 9; Length 894;
 Best Local Similarity 93.7%; Pred. No. 1.3e-123;
 Matches 650; Conservative 0; Mismatches 31; Indels 13; Gaps 3;

QY 1 GGCCTGGAGCTGTGATTCCTCAACCTCCCTCACTCAGGAGGACAGGAGGCTGGGCCCCAGTC 60
 Db 34 GGCCTGGAGCTGTGATTCCTCAACCTCCCTCACTCAGGAGGACAGGAGGCTGGGCCCCAGTC 93

QY 61 TGCCTGACATGATGCTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 Db 94 TGCCTGACATGATGCTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 153

QY 121 CACTGGAACACAGCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Db 154 CACTGGAACACAGCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 213

QY 181 AGGCAAGCCCCCAACTCATATTATGAGTCAGTATCGGCGCTCAGGCGCTTCTTA 240
 Db 214 AGGCAAGCCCCCAACTCATATTATGAGTCAGTATCGGCGCTCAGGCGCTTCTTA 273

QY 241 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 Db 274 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 333

QY 301 TGAGGACGAGGCTGATTTACTGCTGCTCATATACAGAAAGTACTTCTCAT---GTCTT 357
 Db 334 TGAGGACGAGGCTGATTTACTGCTGCTCATATACAGAAAGTACTTCTCAT---GTCTT 393

QY 358 CGGAAGTGGAGCCAGAGTCAACCGTCTAGTCAAGCCCAAGCCCAAGCCCAAGTCACTCT 417
 Db 394 CGGCGAGGAGGACCAAGTCAACCGTCTAGTCAAGCCCAAGCCCAAGCCCAAGTCACTCT 453

QY 418 GTTCCCGCCCTCTGAGAGCTCCCAAGCCCAAGCCCAAGTCTGCTGCTGCTGCTGCTGCTG 477
 Db 454 GTTCCCGCCCTCTGAGAGCTCCCAAGCCCAAGCCCAAGTCTGCTGCTGCTGCTGCTGCTG 504

QY 478 TGACTTCTACCCCGGAGCTGTGACAGTGGCTGAGAGGAGATGGAGCCCGCTCAAGGC 537
 Db 505 TGACTTCTACCCCGGAGCTGTGACAGTGGCTGAGAGGAGATGGAGCCCGCTCAAGGC 564

QY 538 GGGAGTGGAGCCCAAGTCAACCGTCTCAAGAGCAAGCAAGTCAAGTCAAGTCAAGTCAAG 557
 Db 565 GGGAGTGGAGCCCAAGTCAACCGTCTCAAGAGCAAGCAAGTCAAGTCAAGTCAAGTCAAG 624

QY 598 CCTGAGCTGAGCGCCGAGAGCTGGAAGTCCCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAG 657
 Db 625 CCTGAGCTGAGCGCCGAGAGCTGGAAGTCCCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAG 684

QY 658 TGAAGGAGACCGTGTGATGAAGACAGTGTCTCC 691
 Db 685 TGAAGGAGACCGTGTGATGAAGACAGTGTCTCC 717

RESULT 10
 HSIQVL034 756 bp mRNA linear PRI 04-NOV-1994
 LOCUS HSIQVL034 756 bp mRNA linear PRI 04-NOV-1994
 DEFINITION Human rearranged immunoglobulin lambda light chain mRNA.
 ACCESSION X57824
 VERSION X57824.1 GI:33747
 KEYWORDS Ig lambda light chain; immunoglobulin.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 344 to 390)
 AUTHORS Combarato, G. and Klobbeck, H.G.
 TITLE V lambda and J lambda-C lambda gene segments of the human
 immunoglobulin lambda light chain locus are separated by 14 kb and
 rearrange by a deletion mechanism
 JOURNAL Eur. J. Immunol. 21 (6), 1513-1522 (1991)
 MEDLINE 91257162
 PUBMED 1904362

REFERENCE 2 (bases 1 to 756)
 AUTHORS Klobbeck, H.G.
 TITLE Direct Submission
 JOURNAL Submitted (31-JAN-1991) H.G. Klobbeck, Inst fuer Physiologische
 Chemie, Physiologische Biochemie und Zellbiologie der Universitaet
 Muenchen, Shillerstr. 44, 8000 Muenchen 2, Germany
 COMMENT for overlapping sequences see: X51754-55; J00252-54; M15641-42.

FEATURES
 source
 1..756
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /isolate="individual S"
 /db_xref="taxon:9606"
 /chromosome="22"
 /clone="c846"
 /cell_type="B-cell"
 /tissue_type="spleen"
 /clone_lib="phage library cs"
 1..756
 /gene="immunoglobulin lambda light chain"
 1..756
 /gene="immunoglobulin lambda light chain"
 /note="cDNA"
 /evidence=experimental
 1..708
 /gene="immunoglobulin lambda light chain"
 /codon_start=1
 /protein_id="CAA40961.1"
 /db_xref="GI:33748"
 /db_xref="EMBL:CAA40961"
 /translation="MAWTLFLTLTGTGTGSGMAQSALTTPASVSGSPGOSITISCTGT
 SSDIGAYFVSYQYQPGKAPKLIPIVSGPSSGISRFSKSGKNTASLTISGLQAE
 DEADYVCASVRSSTVVGSGSYTVLQGPAASTVLPSPSEELCAKNTLVCLIS
 DEPVATVAMKADSPVRAVETTTTSKSNMNYAASYSLTLPEDKMSRYSQCV
 THESIVETKTPATPBCS"
 1..57
 /gene="immunoglobulin lambda light chain"
 58..388
 /gene="immunoglobulin lambda light chain"
 /note="variable region
 V(lambda)II"
 352..388
 /gene="immunoglobulin lambda light chain"
 /note="J-segment"
 389..708
 /gene="immunoglobulin lambda light chain"
 /note="constant region"

ORIGIN
 Query Match 40.0%; Score 584.6; DB 9; Length 756;
 Best Local Similarity 91.3%; Pred. No. 9.8e-123;
 Matches 631; Conservative 0; Mismatches 59; Indels 1; Gaps 1;

QY 1 GGCCTGGAGCTGTGATTCCTCAACCTCCCTCACTCAGGAGGACAGGAGGCTGGGCCCCAGTC 60
 Db 3 GGCCTGGAGCTGTGATTCCTCAACCTCCCTCACTCAGGAGGACAGGAGGCTGGGCCCCAGTC 62

QY 61 TGCCTGACTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 Db 63 TGCCTGACTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 122

QY 121 CACTGGAACACAGCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Db 123 CACTGGAACACAGCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 182

QY 181 AGGCAAGCCCCCAACTCATATTATGAGTCAGTATCGGCGCTCAGGCGCTTCTTA 240
 Db 189 AGGCAAGCCCCCAACTCATATTATGAGTCAGTATCGGCGCTCAGGCGCTTCTTA 242

QY 241 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 Db 243 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 302


```

QY 301 TGAGGACGAGGCTGATTTATCTGCTCATATACAAAGACTTCTCATCTCTCGG 360
DB 303 TGAGGACGAGGCTGATTTATCTGCTCATATACAAAGACTTCTCATCTCTCGG 362
QY 361 AACTGAGCAACAGTGCACCGCTCTAGTGCAGGCCAAGCCCACTGTCAGTCTGTT 420
DB 363 CGGAGGACACGAGGCTGACCGCTCTAGTGCAGGCCAAGCCCACTGTCAGTCTGTT 422
QY 421 CCCGCCCTCTCTGAGAGCTCCAAAGCAAGGCCCACTAGTGTCTGATCAGTGA 480
DB 423 CCCGCCCTCTCTGAGAGCTCCAAAGCAAGGCCCACTAGTGTCTGATCAGTGA 482
QY 481 CTTTACCCGGGAGGTGAGACAGTGGCCCTGGAAGGCAGATGGCCCGCCGCAAGGGG 540
DB 483 CTTTACCCGGGAGGTGAGACAGTGGCCCTGGAAGGCAGATGGCCCGCCGCAAGGGG 542
QY 541 AGTGAAGACCAACCAACCTCCAAACAGAGCAACCAAGTACGCGGCGCAGACCTACT 600
DB 543 AGTGAAGACCAACCAACCTCCAAACAGAGCAACCAAGTACGCGGCGCAGACCTACT 602
QY 601 GAGCTGAGACCCGAGCAGTGAAGTCCCAAGAGCTACAGCTGCGCAGCTCAGCATGA 660
DB 603 GAGCTGAGACCCGAGCAGTGAAGTCCCAAGAGCTACAGCTGCGCAGCTCAGCATGA 662
QY 661 AGGAGACACCGTGAAGTGAAGACAGTGTCCC 691
DB 663 AGGAGACACCGTGAAGTGAAGTGTGCC 692

RESULT 11
AB064158 810 bp mRNA linear PRI 02-JUL-2002
LOCUS Homo sapiens IGL mRNA for immunoglobulin lambda light chain VLJ
DEFINITION region, partial cds, clone:LI18.
ACCESSION AB064158
VERSION AB064158.1 GI:21669522
KEYWORDS
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 Akahori,Y., Iba,Y., Morino,K., Shinohara,M., Hiroo,Y., Kakita,M.,
Miyata,K. and Kurosawa,Y.
Construction and characterization of antibody libraries: isolation
of therapeutic human antibodies and application to functional
genomics
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 810)
AUTHORS Kurosawa,Y.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-2001) Yoshikazu Kurosawa, Institute for
Comprehensive Medical Science, Fujita Health University;
Kuratsuke-cho, Toyosake 470-1192, Japan
(E-mail:kurosawa@fujita-hu.ac.jp, Tel:81-562-93-9387)
COMMENT Please visit our web site
URL:http://www.fujita-hu.ac.jp/immunity/.
FEATURES
source
1..810
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="LI18"
/clone_11b="ALMS4"
/notes="mixture of tissues:tonsils, umbilical cords,
peripheral blood and bone marrow"
1..810
/gene="IGL"
1..>810
/gene="IGL"
/codon_start=1
/product="immunoglobulin lambda light chain VLJ region"

```

```

sig_peptide
1..66
/gene="IGL"
/notes="pe1b signal peptide"

ORIGIN
Query Match 40.0%; Score 584.4; DB 9; Length 810;
Best Local Similarity 95.0%; Pred. No. 1.1e-122;
Matches 614; Conservative 0; Mismatches 31; Indels 1; Gaps 1;

QY 46 GTCTGGGCGCCAGTCTGCCCTGACTCAGTCTGCTCGGTCTGGGCTCTCTGACAGTC 105
DB 57 GGCATGGCCGCAAGTCTGCCCTGACTCAGCTGCTCGGTCTGGGCTCTCTGACAGTC 116
QY 106 GATCAACATCTCTGACATGGAACAGACAGTACCGTGTGATTAATATGTCCTG 165
DB 117 AGTCAACATCTCTGACATGGAACAGACAGTACCGTGTGATTAATATGTCCTG 176
QY 166 GTACCAACAGACCAAGCAAGCCCAACTCATATTTATGAGTCAATATGAGCC 225
DB 177 GTACCAACAGACCAAGCAAGCCCAACTCATATTTATGAGTCAATATGAGCC 236
QY 226 CTGAGGGGTTTCTAATGCTTCTCTGCTCCAAAGTGTGGCAACAGGCTCCCTACAT 285
DB 237 CTGAGGGGTTTCTAATGCTTCTCTGCTCCAAAGTGTGGCAACAGGCTCCCTACAT 296
QY 286 CTCTGGGCTCAAGGCTGAGGACGAGGCTGATATTAATGCTCTCATATACAAAGTAC 345
DB 297 CTCTGGGCTCAAGGCTGAGGACGAGGCTGATATTAATGCTCTCATATACAAAGTAC 356
QY 346 TTCTCATGCTTCTGGAAGTGGACCAAGTCAACGCTCTAGTCAAGCCCAAGCC 405
DB 357 CACTGGGCTTCTGGAAGTGGACCAAGTCAACGCTCTAGTCAAGCCCAAGCC 416
QY 406 CACTGGGCTTCTGGAAGTGGACCAAGTCAACGCTCTAGTCAAGCCCAAGCC 465
DB 417 CACTGGGCTTCTGGAAGTGGACCAAGTCAACGCTCTAGTCAAGCCCAAGCC 476
QY 466 GTCTGATCAGTGAATTTACCCGGAGCTGTGACAGTGGCTGAAAGCAGATGCGAG 525
DB 477 GTCTGATCAGTGAATTTACCCGGAGCTGTGACAGTGGCTGAAAGCAGATGCGAG 536
QY 526 CCCGCTCAAGGGGAGTGGAGACCAACCAACCTCCAAACAGACCAACCAAGTACGC 585
DB 537 CCCGCTCAAGGGGAGTGGAGACCAACCAACCTCCAAACAGACCAACCAAGTACGC 596
QY 586 GGCACAGACTACCTGAGCCTGACGCCCGAGAGAGTGAAGTCCACAGAGCTACAGCTG 645
DB 597 GGCACAGACTACCTGAGCCTGACGCCCGAGAGAGTGAAGTCCACAGAGCTACAGCTG 656
QY 646 CCAAGTCAAGCAATGAAGGAGACCGTGAATGAAGACAGTGTCCC 691
DB 657 CCAAGTCAAGCAATGAAGGAGACCGTGAATGAAGACAGTGTCCC 701

RESULT 12
BC015833 915 bp mRNA linear PRI 19-NOV-2003
LOCUS Homo sapiens cDNA clone MGC:27152 IMAGE:4691630, complete cds.
ACCESSION BC015833
VERSION BC015833.1 GI:16198374
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 915)

```


rearrange by a deletion mechanism
Eur. J. Immunol. 21 (6), 1513-1522 (1991)
MEDLINE
91257162
PUBMED
1904362
2 (bases 1 to 783)
Klobeck, H.G.
AUTHORS
Direct Submission
JOURNAL
Submitted (31-JUN-1991) H.G. Klobeck, Inst fuer Physiologische
Chemie, Physiologische Biochemie und Zellbiologie der Universitaet
Muenchen, Shillerstr. 44, 8000 Muenchen 2, Germany
COMMENT
for overlapping sequences see: X51754-55; J00252-54; M15641-42.
FEATURES
source
1..783
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="C2132"
/cell_line="GM2132"
/cell_type="B-cell"
/clone_lib="phage library c2132"
1..783
/gene="immunoglobulin lambda light chain"
1..793
/gene="immunoglobulin lambda light chain"
/note="cDNA"
/evidence="experimental"
23..736
/gene="immunoglobulin lambda light chain"
/codon_start=1
/protein_id="CAA0939.1"
/db_xref="GI:33700"
/db_xref="RENTREMBL:CAA0939"
/translation="NAMALLTLTLTGSGMSAQSALNPASVNSPGQILITSGTP
SSDIDGYQYISWYQGHKPKAKPLIITYDVXKRSQSVSNRSGSKSNTSLITSGIQA
DEADYICSSIRSGATFEVFGGATKVTVPKAPASVTLFPSEELQAKKATIVCL
ISDFYPGAVTAMKADSSPVKAGVETTPSKQSNKRYAASVYLSLTPQWMSHRSYSC
QVTHSGSTVEKTVAPRECS"
23..79
/gene="immunoglobulin lambda light chain"
80..416
/gene="immunoglobulin lambda light chain"
/note="variable region
V(lambda)II"
380..416
/gene="immunoglobulin lambda light chain"
/note="J-segment"
417..736
/gene="immunoglobulin lambda light chain"
/note="constant region"
ORIGIN
Query Match 39.1%; Score 570.2; DB 9; Length 783;
Best Local Similarity 90.7%; Pred. No. 1.9e-119;
Matches 632; Conservative 0; Mismatches 58; Indels 7; Gaps 2;
QY 1 GGCCTGGGCTCTGCTATTCCTCAACCTCTCACTAGGCAAGGAGTCTGGGCCCAAGTC 60
DB 25 GGCCTGGGCTCTGCTCTCTCTCAACCTCTCACTAGGCAAGGAGTCTGGGCCCAAGTC 84
QY 61 TGCCCTGACTCACTGACCTGCGGTGCTGGGCTCTCTGAGCAGTGATCAACATCTCTG 120
DB 85 TGCCCTGACTCACTGACCTGCGGTGCTGGGCTCTCTGAGCAGTGATCAACATCTCTG 144
QY 121 CACTGAAACCAAGAGTCAAGTGGTGTATATATATATATATATATATATATATATATAT 180
DB 145 CACTGAAACCAAGAGTCAAGTGGTGTATATATATATATATATATATATATATATATAT 204
QY 181 AGGCAAG 240
DB 205 AGGCAAG 264
QY 241 TGCTTCTCTGCTCAAGTCTGGCAACAGGCTCTCTGACCAATCTCTGGGCTCCAGGC 300

DB 265 TGCTTCTCTGCTCAAGTCTGGCAACAGGCTCTCTGACCAATCTCTGGGCTCCAGGC 324
QY 301 TGAAG 354
DB 325 TGAAG 384
QY 355 CTTCGAG 414
DB 385 GTTCGAG 444
QY 415 TCTGTCCTCCGCTCTCTGAG 474
DB 445 TCTGTCCTCCGCTCTCTGAG 504
QY 475 CAGTGAATTCACCCGAG 534
DB 505 AAGTGAATTCACCCGAG 564
QY 535 GCGGAG 594
DB 565 GCGGAG 624
QY 595 CTACCTGAG 654
DB 625 CTACCTGAG 684
QY 655 GCATGAAG 691
DB 685 GCATGAAG 720
RESULT 14
AB064155 806 bp mRNA linear PRI 02-JUN-2002
LOCUS
DEFINITION Homo sapiens IGL mRNA for immunoglobulin lambda light chain VJJ
region, partial cds, clone:115.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (25-JUN-2001) Yoshikazu Kurosawa, Institute for
Comprehensive Medical Science, Fujita Health University,
Kutsukake-cho, Toyoake 470-1192, Japan
(E-mail:kurosawa@fujita-hu.ac.jp, Tel:81-562-93-9387)
Please visit our web site
URL: http://www.fujita-hu.ac.jp/immunity/
FEATURES
source
1..806
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="115"
/clone_lib="AIMS4"
/note="mixture of tissues: tonsils, umbilical cords,
peripheral blood and bone marrow"
1..806
/gene="IGL"
1..>806


```
Db 357 CTCTTGAGTGTTCGGCGAGAGGACCAAGCTGACCGTCTAGGTCAGCCCAAGGCTGCCCC 416
QY 406 CACTGTCACTCTGTGTTCCCGCCCTCTCTGAGAGCTCCAAGCCAAAGGCCACACTAGT 465
Db 417 CTGGTCACTGTGTTCGCCCTCTCTGAGGAGCTTCAAGCCAAAGGCCACACTGAT 476
QY 466 GTGTCTGATCAGTGACTTCTTACCCGAGAGCTGTGACAGTGCCTGAAAGGCAGATGCGAG 525
Db 477 GTGTCTCATAGTGACTTCTTACCCGAGAGCTGTGACAGTGCCTGAAAGGCAGATGCGAG 536
QY 526 CCCCGTCAAGGCGGAGGTGAGACCAACCAACCTTCCAAACAGAGCAACACAGTACGC 585
Db 537 CCCCGTCAAGGCGGAGGTGAGACCAACCAACCTTCCAAACAGAGCAACACAGTACGC 596
QY 586 GGCACAGAGCTACTGAGCCTGACGCCGAGCAGTGAAGTCCACAGAAAGCTACAGCTG 645
Db 597 GGCACAGAGCTACTGAGCCTGAGCCTGAGCAGTGAAGTCCACAGAAAGCTACAGCTG 656
QY 646 CCAAGTCAAGCAGTGAAGGAGCAACCGTGAATGAAGCAAGTGTGCC 691
Db 657 CCAAGTCAAGCAGTGAAGGAGCAACCGTGAATGAAGCAAGTGTGCC 701
```

Search completed: April 26, 2004, 18:46:55
Job time : 5606.61 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 14:04:56 ; Search time 555.73 Seconds
(without alignments)
11160.755 Million cell updates/sec

Title: US-10-076-747-8

Perfect score: 1460

Sequence: 1 ggcctggcctctgtatcc.....ggcagcagagccacgcagct 1460

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 3373663 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002s:*
7: Geneseq2003as:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1459	99.9	1460	7	ABX92194 Human ova
2	809.6	55.5	957	7	ABX92193 Human ova
3	616.8	42.2	1636	5	ABV22585 Human pro
4	616.8	42.2	1636	5	ABV28405 Human pro
5	607	41.6	891	4	AAC65528 Human imm
6	597.4	40.9	902	2	AAQ35100 Antibody
7	597.4	40.9	1480	8	ADAI1028 Human CDN
8	595.8	40.8	763	5	AA883485 DNA encod
9	589.8	40.4	756	5	AA883477 DNA encod
10	584.6	40.0	1636	5	ABV22585 Human pro
11	584.6	40.0	1636	5	ABV28405 Human pro
12	581.2	39.8	2667	9	AA851189 Human rea
13	581.2	39.8	2667	9	AA851189 Human rea
14	578.2	39.6	889	5	AA851186 Human (hy
15	575.2	39.4	783	5	AA883483 DNA encod
16	570.2	39.1	866	5	AA887037 DNA encod
17	564.8	38.7	876	5	AA883478 DNA encod
18	562.5	38.5	811	9	ADP07015 Novel cod
19	558.6	38.1	726	7	ABX12863 DNA encod
20	529.4	36.3	2112	6	ABQ75369 Human lun
21	511.8	35.1	777	5	AA877072 DNA encod
22	508	34.8	935	4	AAC66525 Human imm
23	498.4	34.1			

24	497.2	34.1	605	6	ABK29738 Colon ade
25	488.6	33.5	9532	7	AB237360 MALTA3 nu
26	482	33.0	870	6	AA150812 Human can
27	480.6	32.9	543	6	ABK38932 CDNA enco
28	480.6	32.9	543	7	ACA11261 Human lun
29	480.6	32.9	543	7	ACA02447 Lung can
30	480.4	32.9	884	2	AAQ03609 Sequence
31	477.4	32.7	1845	6	AB149526 Plasmid s
32	473.4	32.4	930	4	AA822541 Human CDN
33	473	32.4	848	6	AA859183 Human Ig
34	468.2	32.1	915	6	AB165478 Lung can
35	468.2	32.1	915	6	ABK64815 Human ben
36	468.2	32.1	915	6	ABN97248 Gene #374
37	468.2	32.1	915	7	ACA64802 Human IG
38	468.2	32.1	964	6	ABQ93374 Human CDN
39	466.6	32.0	762	5	AA884209 Plasmid G
40	465.2	31.9	5679	5	AA884207 Plasmid G
41	463.4	31.7	951	5	AA890539 DNA encod
42	462.8	31.7	888	5	AA859184 Human Ig
43	459.6	31.5	810	5	AA887270 DNA encod
44	458	31.4	768	2	AA806954 Monoclon
45	458	31.4	768	2	AA806953 Monoclon

ALIGNMENTS

RESULT 1	ABX92194	Standard, cDNA, 1460 BP.
XX	ABX92194;	
AC	ABX92194;	
XX	08-MAY-2003 (first entry)	
DT	08-MAY-2003 (first entry)	
XX	Human ovarian specific nucleic acid DEX0310_8.	
XX	Human; gene; ss; ovarian specific nucleic acid; OSNA; ovarian cancer;	
KW	non-cancerous ovarian disease; gene therapy; vaccine; cyostatic;	
KM	gynaecological.	
XX		
OS	Homo sapiens.	
XX		
XX	WO200292785-A2.	
PN	21-NOV-2002.	
PD	13-FEB-2002; 2002W0-US022271.	
PF	13-FEB-2001; 2001US-0268290P.	
XX	13-FEB-2001; 2001US-0268834P.	
PR	(DIAD-) DIADEXUS INC.	
PA	Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferkey R;	
PI	Sun Y, Liu C;	
XX	WPI; 2003-120677/1.	
DR	P-PSDB; ABU61023.	
XX	New isolated OSNA nucleic acid and encoded polypeptide, useful for	
PT	identifying, diagnosing, monitoring, staging, imaging and treating	
PT	ovarian cancer and non-cancerous diseases in ovarian tissues.	
PT	Claim 1; Page 146; 224pp; English.	
PS	The invention relates to a new isolated nucleic acid termed ovarian	
XX	specific nucleic acid (OSNA) comprising: (a) a nucleic acid molecule that	
CC	encodes any of 53 fully defined protein sequences appearing as ABU61018-	
CC	ABU61070 (termed ovarian specific proteins OSP); (b) any of 76 fully	
CC	defined nucleotide sequences appearing as ABX92187-ABX92262; or (c) a	
CC	sequence having at least 60% sequence identity to the nucleic acid	
CC	molecule of (a) or (b). Also included are a method for determining the	

CC presence of an ovary specific nucleic acid (OSNA) in a sample, a vector
 CC comprising an OSNA, a host cell comprising the vector, an isolated OSP
 CC polypeptide, an anti-OSP antibody or fragment, a method for determining
 CC the presence of an ovary specific protein in a sample and a vaccine
 CC comprising an OSP or OSNA. The methods and compositions of the present
 CC invention are useful for identifying, diagnosing, monitoring, staging,
 CC imaging and treating ovarian cancer and non-cancerous disease in ovary
 CC tissue. The present sequence is an OSNA of the invention
 XX

SQ Sequence 1460 BP; 351 A; 466 C; 401 G; 241 T; 0 U; 1 Other;

Query Match 99.9%; Score 1459; DB 7; Length 1460;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1460; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGCCTGGGCTGCTATTCCTCACTCACTGAGGCAAGGCTCTGGGCGCCAGTC 60
DB 1 GGCCTGGGCTGCTATTCCTCACTCACTGAGGCAAGGCTCTGGGCGCCAGTC 60
QY 61 TGCCCTGACTAGTCTGCTCCGTCTGAGTCTCTGAGAGTCAACATCTCTG 120
DB 61 TGCCCTGACTAGTCTGCTCCGTCTGAGTCTCTGAGAGTCAACATCTCTG 120
QY 121 CACTGGAACAGAGTCAAGTGTGTGTATTAATATCTCTCTGTAACAAGACC 180
DB 121 CACTGGAACAGAGTCAAGTGTGTGTATTAATATCTCTCTGTAACAAGACC 180
QY 181 AGGCAAGGCCCCCAATCATATTTATGAGTCAATGAGGCTCAGAGGCTTTTAA 240
DB 181 AGGCAAGGCCCCCAATCATATTTATGAGTCAATGAGGCTCAGAGGCTTTTAA 240
QY 241 TGGCTTCTCTGCTCAAGTCTGAGCAAGGCTCTCTGACATCTCTGAGGCTCAAGC 300
DB 241 TGGCTTCTCTGCTCAAGTCTGAGCAAGGCTCTCTGACATCTCTGAGGCTCAAGC 300
QY 301 TGAAGACAGAGGTGATTAATGCTGCTCAATTAAGAGTCTCTCTCTCTCTG 360
DB 301 TGAAGACAGAGGTGATTAATGCTGCTCAATTAAGAGTCTCTCTCTCTCTG 360
QY 361 AACTGGAGCAAGTCAAGTCTCTGAGTCAAGGCAAGGCTCAAGGCTCAAGTCT 420
DB 361 AACTGGAGCAAGTCAAGTCTCTGAGTCAAGGCAAGGCTCAAGGCTCAAGTCT 420
QY 421 CCGGCTCTCTCTGAGAGTCTCAAGCAAGGCTCAAGTCTCTCTCTCTCTCTG 480
DB 421 CCGGCTCTCTCTGAGAGTCTCAAGCAAGGCTCAAGTCTCTCTCTCTCTCTG 480
QY 481 CTTCTACCGGAGCTGTAAGTCAAGTCTGAGGCAAGGCTCAAGGCTCAAGGCT 540
DB 481 CTTCTACCGGAGCTGTAAGTCAAGTCTGAGGCAAGGCTCAAGGCTCAAGGCT 540
QY 541 AGTGAAGACCAACCACTCTCAACAGAGCAAGTCAAGGCTCAAGGCTCAAGTCT 600
DB 541 AGTGAAGACCAACCACTCTCAACAGAGCAAGTCAAGGCTCAAGGCTCAAGTCT 600
QY 601 GAGGCTGAGGCGGAGAGTGAAGTCCCAAGAGTCAAGGCTCAAGGCTCAAGTGA 660
DB 601 GAGGCTGAGGCGGAGAGTGAAGTCCCAAGAGTCAAGGCTCAAGGCTCAAGTGA 660
QY 661 AGGAGACACCGTGAAGAGTGTCTCTCTCAAGAAATGTTCTATAGTCTCTAAC 720
DB 661 AGGAGACACCGTGAAGAGTGTCTCTCTCAAGAAATGTTCTATAGTCTCTAAC 720
QY 721 AGGCTCAGGCGGCACTACCAATGAGGAGACTAGCAAGTCAAGGCTCAAGGAGG 780
DB 721 AGGCTCAGGCGGCACTACCAATGAGGAGACTAGCAAGTCAAGGCTCAAGGAGG 780
QY 781 GGTCTCTCCATCAACAACAAGAGGAGTCAAGGCTCTATATCTCTCTCTCTGAC 840
DB 781 GGTCTCTCCATCAACAACAAGAGGAGTCAAGGCTCTATATCTCTCTCTCTGAC 840
QY 841 GTCTATTAAGAGGCTCAAGAAATTAAGATTTCTGACTTGTGAATGTTCTCTACATCA 900
DB 841 GTCTATTAAGAGGCTCAAGAAATTAAGATTTCTGACTTGTGAATGTTCTCTACATCA 900

```

```

DB 841 GTCTATTAAGAGGCTCAAGAAATTAAGATTTCTGACTTGTGAATGTTCTCTACATCA 900
QY 901 TTAACAAAGACATATCTCACTACAGAGTGTGATGATCAACCGGCTCGTGGGCGAAT 960
DB 901 TTAACAAAGACATATCTCACTACAGAGTGTGATGATCAACCGGCTCGTGGGCGAAT 960
QY 961 GGCACCTGCTCTGCTGCTCAATGATGATCTCAAGTGAACAACAGGCTGTGGCGCAC 1020
DB 961 GGCACCTGCTCTGCTGCTCAATGATGATCTCAAGTGAACAACAGGCTGTGGCGCAC 1020
QY 1021 ANCTATGAGTGAATATTCGCGGAGTCAAGACCATGATTTCCCGGCTGATCG 1080
DB 1021 ANCTATGAGTGAATATTCGCGGAGTCAAGACCATGATTTCCCGGCTGATCG 1080
QY 1081 AATATGCGGTACGCGGCAATCAATTCGCGGAGCTACAGACTTAAAGTTGACCGC 1140
DB 1081 AATATGCGGTACGCGGCAATCAATTCGCGGAGCTACAGACTTAAAGTTGACCGC 1140
QY 1141 GCAGCGGCGGAAAGCAAGGCTTTGACAGGAATGCCAAACAAGAGGCGGACAGC 1200
DB 1141 GCAGCGGCGGAAAGCAAGGCTTTGACAGGAATGCCAAACAAGAGGCGGACAGC 1200
QY 1201 GCGGACCCCGGCGGAGAGGCCCCACTGCGGAGGCGAGGAGGAGCAAGTACAGGA 1260
DB 1201 GCGGACCCCGGCGGAGAGGCCCCACTGCGGAGGCGAGGAGGAGCAAGTACAGGA 1260
QY 1261 CCGGAGCCACAGCGCGCCCTGTGATGAGACGAGACGCGGAGGAGGAGCGAGCCGAA 1320
DB 1261 CCGGAGCCACAGCGCGCCCTGTGATGAGACGAGACGCGGAGGAGGAGCGAGCCGAA 1320
QY 1321 GCAAGCGCGCAAGAGTCAAGCGGCAAGGCTCAAGGCTCTGAGGCTCTGAGGCTCGG 1380
DB 1321 GCAAGCGCGCAAGAGTCAAGCGGCAAGGCTCAAGGCTCTGAGGCTCTGAGGCTCGG 1380
QY 1381 CAGCGCGAGAGGCGGCGGCGGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
DB 1381 CAGCGCGAGAGGCGGCGGCGGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
QY 1441 GGCAGCAGAGCCACCGAGCT 1460
DB 1441 GGCAGCAGAGCCACCGAGCT 1460

```

RESULT 2

ABX92193

ID ABX92193 standard; cDNA; 957 BP.

XX ABX92193;

DT 08-MAY-2003 (first entry)

XX Human ovarian specific nucleic acid DEX010_7.

XX Human; gene; ss; ovarian specific nucleic acid; OSNA; ovarian cancer;

KW non-cancerous ovarian disease; gene therapy; vaccine; cytostatic;

KW gynaecological.

OS Homo sapiens.

XX WO200292785-A2.

XX 21-NOV-2002.

XX 13-FEB-2002; 2002WO-US022271.

XX 13-FEB-2001; 2001US-0268290P.

XX 15-FEB-2001; 2001US-0268834P.

XX (DIAD-) DIADEXUS INC.

XX Salceda S, Macina RA, Hu P, Recipon H, Karra K, Caferkey R;

PI Sun Y, Liu C;

XX

DR MPI: 2003-120677/11.
 DR P-P8DB; ABU61022.
 XX New isolated OSNA nucleic acid and encoded polypeptide, useful for
 PT identifying, diagnosing, monitoring, staging, imaging and treating
 PT ovarian cancer and non-cancerous diseases in ovarian tissues.
 XX
 PS Claim 1; Page 145; 224pp; English.

XX The invention relates to a new isolated nucleic acid, termed ovarian
 CC specific nucleic acid (OSNA) comprising: (a) a nucleic acid molecule that
 CC encodes any of 53 fully defined protein sequences appearing as ABU61018-
 CC ABU61070 (termed ovarian specific proteins, OSP); (b) any of 76 fully
 CC defined nucleotide sequences appearing as ABX92187-ABX92262; or (c) a
 CC sequence having at least 60% sequence identity to the nucleic acid
 CC molecule of (a) or (b). Also included are a method for determining the
 CC presence of an ovary specific nucleic acid (OSNA) in a sample, a vector
 CC comprising an OSNA, a host cell comprising the vector, an isolated OSP
 CC polypeptide, an anti-OSP antibody or fragment, a method for determining
 CC the presence of an ovary specific protein in a sample and a vaccine
 CC comprising an OSP or OSNA. The methods and compositions of the present
 CC invention are useful for identifying, diagnosing, monitoring, staging,
 CC imaging and treating ovarian cancer and non-cancerous disease in ovary
 CC tissue. The present sequence is an OSNA of the invention
 CC
 CC

SQ Sequence 957 BP; 238 A; 290 C; 288 G; 140 T; 0 U; 1 Other;

Query Match 55.5%; Score 809.6; DB 7; Length 957;
 Best Local Similarity 96.5%; Pred. No. 1.3e-180;
 Matches 864; Conservative 0; Mismatches 19; Indels 12; Gaps 3;

QY 578 AAGTACGGCGCAGCAGCTACCTGAGCCCTGACGCC-----GAGCAGTGAAGTCCACAG 633
 DB 63 AGGTACCGCGTCAGAGCTATCTGAGCCTGACGCCCTGATGACATGAGAGTCCACAG 122
 QY 634 AAGTACAGCTGCCAGCTAC-----GCATGAAGGAG-CACCTGTGATGAAGACGT 685
 DB 123 AAGTACAGCCTGACAGAGTCTTAGCTGATGAAGGAGACACCGTGTGAAGACGT 182
 QY 686 GATCCCTACAGAAATGTTCTATAGTCTCTTAACAGCCTAGCCCACTACCAATGAG 745
 DB 183 GGTCCCTACAGAAATGTTCTATAGTCTCTTAACAGCCTAGCCCACTACCAATGAG 242
 QY 746 GAGACTACAGCTGACAGGATCCCAAGGAGAGGGGCTCTTCATCAACACACACAG 805
 DB 243 GAGACTACAGCTGACAGGATCCCAAGGAGAGGGGCTCTTCATCAACACACACAG 302
 QY 806 GCGGAGTAAAGCCCTTATCATCGGCAATGTCGACGTCATGAAGGCTCAATAA 865
 DB 303 GCGGAGTAAAGCCCTTATCATCGGCAATGTCGACGTCATGAAGGCTCAATAA 362
 QY 866 GATATTCGACTTGGTGAATGTCCTACATACATACAAAGACATATCACTACA 925
 DB 363 GATATTCGACTTGGTGAATGTCCTACATACATACAAAGACATATCACTACA 422
 QY 926 CGGAGTCATGATACCGGTCGCGGCGGAAATGCACTCCCTCTCGGCGCTCAAT 985
 DB 423 CGGAGTCATGATACCGGTCGCGGCGGAAATGCACTCCCTCTCGGCGCTCAAT 482
 QY 986 GATGACTATGATACACCGGTCGCGGCGGCAANTATGATGATATTCGCCGGG 1045
 DB 483 GATGACTATGATACACCGGTCGCGGCGGCAANTATGATGATATTCGCCGGG 542
 QY 1046 ATGACAGACATGATTTCCCGCGTGGCTGATGATATGCGGTAGCGGCATCAATT 1105
 DB 543 ATGACAGACATGATTTCCCGCGTGGCTGATGATATGCGGTAGCGGCATCAATT 602
 QY 1106 CGCCCGGAGGCTACAGACCTTAAAGTTGACCGGCGAGCGCGGAGAAACAGGCTTTCG 1165
 DB 603 CGCCCGGAGGCTACAGACCTTAAAGTTGACCGGCGAGCGCGCGGAGAAACAGGCTTTCG 662
 QY 1166 ACGGAATGCCAAACACAGAGGCGCGACAGCCGGCAGCGCTCCGGCGGAGAGCCCCA 1225

DB 663 ACGGAATGCCAAACACAGAGGGCCGACAGCCGGCAGGAGACCCCGGCGAGAGCCCCA 722
 QY 1226 CTGCGGACAGGCGGAGCGGAGGACGATACGAGACGCGGACACACCGCGCCGTGCA 1285
 DB 723 CTGCGGACAGGCGGAGCGGAGGACGATACGAGACGCGGACACACCGCGCCGTGCA 782
 QY 1286 TGAGACGAGACGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 1345
 DB 783 TGAGACGAGACGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 842
 QY 1346 AGCCACAGCGGCTTGAAGCCCTGCGGACGCGCGGACGCGGAGGCGGCGGAGCA 1405
 DB 843 AGCCACAGCGGCTTGAAGCCCTGCGGACGCGCGGACGCGGAGGCGGCGGAGCA 902
 QY 1406 CGGCGAGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACGT 1460
 DB 903 CGGCGAGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACGT 957

RESULT 3

ID ABV22585
 XX ABV22585 strand; cDNA; 1636 BP.

AC ABV22585;
 XX

DT 13-SEP-2002 (first entry)
 XX

DE Human prostate expression marker cDNA 22576.
 XX

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX

OS Homo sapiens.
 XX

PN MO200160860-A2.
 XX

PD 23-AUG-2001.
 XX

PF 20-FEB-2001; 2001WO-US005171.
 XX

PR 17-FEB-2000; 2000US-0189319P.
 XX

PR 16-MAR-2000; 2000US-0189862P.
 XX

PR 25-MAY-2000; 2000US-0207454P.
 XX

PR 09-JUN-2000; 2000US-0211314P.
 XX

PR 18-JUL-2000; 2000US-0219007P.
 XX

PR 13-DEC-2000; 2000US-0255281P.
 XX

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;
 XX

DR MPI; 2001-662795/76.
 XX

PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX

PS Claim 1; Page 3949; 11750pp; English.
 XX

XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 XX Sequence 1636 BP; 345 A; 470 C; 479 G; 342 T; 0 U; 0 Other;

Query Match 42.2%; Score 616.8; DB 5; Length 1636;
 Best Local Similarity 95.7%; Pred. No. 3.3e-135;
 Matches 668; Conservative 0; Mismatches 22; Indels 8; Gaps 3;

1 GGCCTGGGCTCTGCTATTCCTCAACCTCTCACTCAGGAGCAGAGGCTCTGGGCCAGTC 60
 62 GGCCTGGGCTCTGCTATTCCTCAACCTCTCACTCAGGAGCAGAGGCTCTGGGCCAGTC 121
 61 TGCCTGACTGAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 120
 122 TGCCTGACTGAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 181
 121 CACTGGAACGACGATCAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 179
 182 CACTGGAACGACGATCAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 241
 180 CAGGCAAGCCCCCAAACTCATTTATGAGGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 239
 242 CAGGCAAGCCCCCAAACTCATTTATGAGGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 301
 240 ATCCCTTCTCTGCTCTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 239
 302 ATCCCTTCTCTGCTCTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 361
 300 CTGAGGACGAGGCTGATTTATTTACTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 353
 362 CTGAGGACGAGGCTGATTTATTTACTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 421
 354 TCTTGGAACTGGGACCAAGGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 413
 422 TCTTGGAACTGGGACCAAGGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 481
 414 CTCTGTTCCGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 473
 482 CTCTGTTCCGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 541
 474 TCACTGACTTCTTACCCGCGGAGCTGTGACAGTGGCTGGAAGGACGATGAGCCCGTCA 533
 542 TCACTGACTTCTTACCCGCGGAGCTGTGACAGTGGCTGGAAGGACGATGAGCCCGTCA 601
 534 AGGCGGAGTGGAGACCAACCAACCTTCCAAACGAGACCAACCAACCAACCAACCA 593
 602 AGGCGGAGTGGAGACCAACCAACCTTCCAAACGAGACCAACCAACCAACCAACCA 661
 594 GCTACTGAGCTGAGCGCCGAGCAGTGAAGTCCCAAGAGTACGCTGCGCAGGTCA 653
 662 GCTACTGAGCTGAGCGCCGAGCAGTGAAGTCCCAAGAGTACGCTGCGCAGGTCA 721
 654 CGCATGAAGGAGCAACCGTGAATGAAGACAGTGTCCC 691
 722 CGCATGAAGGAGCAACCGTGAATGAAGACAGTGTCCC 758

RESULT 4
 ID ABV28405
 ID ABV28405 standard; cDNA; 1636 BP.
 XX
 AC ABV28405;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 28396.
 XX
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX pharomacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX

PF 20-FEB-2001; 2001WO-US005171.
 XX
 XX 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR MPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 5921; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Table 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC / (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 1636 BP; 345 A; 470 C; 479 G; 342 T; 0 U; 0 Other;

Query Match 42.2%; Score 616.8; DB 5; Length 1636;
 Best Local Similarity 95.7%; Pred. No. 3.3e-135;
 Matches 668; Conservative 0; Mismatches 22; Indels 8; Gaps 3;

1 GGCCTGGGCTCTGCTATTCCTCAACCTCTCACTCAGGAGCAGAGGCTCTGGGCCAGTC 60
 62 GGCCTGGGCTCTGCTATTCCTCAACCTCTCACTCAGGAGCAGAGGCTCTGGGCCAGTC 121
 61 TGCCTGACTGAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 120
 122 TGCCTGACTGAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 181
 121 CACTGGAACGACGATCAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 179
 182 CACTGGAACGACGATCAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 241
 180 CAGGCAAGCCCCCAAACTCATTTATGAGGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 239
 242 CAGGCAAGCCCCCAAACTCATTTATGAGGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 301
 240 ATCCCTTCTCTGCTCTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 239
 302 ATCCCTTCTCTGCTCTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 361
 300 CTGAGGACGAGGCTGATTTATTTACTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 353
 362 CTGAGGACGAGGCTGATTTATTTACTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 421
 354 TCTTGGAACTGGGACCAAGGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 413
 422 TCTTGGAACTGGGACCAAGGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 481
 414 CTCTGTTCCGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 473
 482 CTCTGTTCCGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 541
 474 TCACTGACTTCTTACCCGCGGAGCTGTGACAGTGGCTGGAAGGACGATGAGCCCGTCA 533

QY	181	AGGCAAGACCCCAAACTCATTTAATGAGGTCAAGTATCGGCTCAGGGTTTCAA	24
Db	214	AGGCAAGACCCCAAAATCATTTAATGAGGTCAAGTATCGGCTCAGGGTTTCAA	272
QY	241	TGCTTCTCTGCTCCAAAGTCTGGCAACAAGGCTCTCTGACATCTTGGGCTCCAGGC	300
Db	274	TGCTTCTCTGCTCCAAAGTCTGGCAACAAGGCTCTCTGACATCTTGGGCTCCAGGC	333
QY	301	TGAGGACGAGGCTGATTATTACTGCTGCTCATATACAGAAAGTACTTCTCATGTCTT	360
Db	334	TGAGGACGAGGCTGATTATTACTGCTGCTCATATACAGAAAGTACTTCTCATGTCTT	393
QY	361	AACGTGGACCAAGGTCAAGCTCTTAAAGTCAAGGCCAAAGCCACTGTCACTCTGT	422
Db	394	CGAGAGGACCAAACTACAGCTCTTAAAGTCAAGGCCAAAGCCACTGTCACTCTGT	453
QY	421	CCGGCCCTCTCTGAGAGAGCTCCAAAGCAACAAGGCCACACATGAGTGTGATCAGTGA	480
Db	454	CCGGCCCTCTCTGAGAGAGCTCCAAAGCAACAAGGCCACACATGAGTGTGATCAGTGA	511
QY	481	CTTCTACCGGGAGCTGTGACAGTGGCTTGAAGGACAGTGGACGCCCTCAAAGGCGGG	540
Db	514	CTTCTACCGGGAGCTGTGACAGTGGCTTGAAGGACAGTGGACGCCCTCAAAGGCGGG	573
QY	541	AGTGGAGACCAACCAACCTCTCCAAACAGACAAACAAGTACGGGGCAGACGTAAGCT	600
Db	574	AGTGGAGACCAACCAACCTCTCCAAACAGACAAACAAGTACGGGGCAGACGTAAGCT	633
QY	601	GACCTCGAGCGCCGAGCAGTGGAAAGTCCACAGAAAGTACAGCTGCGACGATGA	666
Db	634	GACCTCGAGCGCCGAGCAGTGGAAAGTCCACAGAAAGTACAGCTGCGACGATGA	699
QY	661	AGGAGACCAAGGTGGATGAAAGACAGTGGTCCC	691
Db	694	AGGAGACCAAGGTGGAAAGTGGAGTGGTCCC	723
RESULT 7			
ADAl1028/C			
ID	ADAl1028	standard, cDNA, 1480 BP.	
XX	ADAl1028/		
XX	AC		
DT	06-NOV-2003	(first entry)	
DE	Human cDNA differentially expressed in colon cancer #101.		
XX	ss; differential expression; colon cancer; cancer; human.		
KW	Homo sapiens.		
XX	OS		
XX	PN	US2002160382-A1.	
PD	31-OCT-2002.		
XX	PF	11-OCT-2001; 2001US-00981351.	
XX	PR	11-OCT-2000; 2000US-0239841P.	
XX	PA	(LASEK/) LASEK A W.	
XX	PI	(JONE/) JONES D A.	
XX	Lasek AW, Jones DA;		
XX	WPI; 2003-265756/26.		
XX	New combination comprising cDNAs that are differentially expressed in		
PT	colon disorder, useful for diagnosing, treating, staging or monitoring		
PT	treatment for colon cancers.		
XX	Claim 1; SEQ ID NO 146; 231p; English.		
XX	The invention relates to a combination comprising cDNAs that are		
CC			

CC differentially expressed in colon disorder. The methods and compositions
 CC of the present invention are useful for diagnosing, treating, staging or
 CC monitoring treatment for colon cancer. They are also useful in high
 CC throughput methods for using cDNAs to detect differential expression of
 CC nucleic acids in a sample, screening molecules or compounds to identify a
 CC ligand which specifically binds a cDNA and using a protein to screen
 CC molecules or compounds to identify at least one ligand which specifically
 CC binds the protein. The present sequence represents a human cDNA
 CC differentially expressed in colon cancer.

XX Sequence 1480 BP; 312 A; 412 C; 454 G; 299 T; 0 U; 3 Other;

Query Match 40.8%; Score 597.4; DB 8; Length 1480;

Best Local Similarity 93.1%; Pred. No. 1.2e-130;
 Matches 649; Conservative 0; Mismatches 41; Indels 7; Gaps 2;

QY 1 GGCCTGGGCTCTGCTATTCTTCCATCCCTCTCCTCAGGAGGAGGCTGGGCGCCAGTGC 60
 DB 1441 GGCCTGGGCTCTGCTATTCTTCCATCCCTCTCCTCAGGAGGAGGCTGGGCGCCAGTGC 1382
 QY TGGCCCTGACTAGTGTGCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 DB 1381 TGGCCCTGACTAGTGTGCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1322
 QY 121 CACTGGAAACGACATCGTGGTGGTATTATGATGCTCTCTGCTGCTGCTGCTGCTGCTGCTG 180
 DB 1321 CACTGGAAACGACATCGTGGTGGTATTATGATGCTCTCTGCTGCTGCTGCTGCTGCTGCTG 1262
 QY 181 AGGCAAAAGCCCCCAACTCATCTATTATGAGTCAATGATGAGGCTGCTGCTGCTGCTGCTGCT 240
 DB 1261 AGGCAAAAGCCCCCAACTCATCTATTATGAGTCAATGATGAGGCTGCTGCTGCTGCTGCTGCT 1202
 QY 241 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 DB 1201 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1142
 QY 301 TGAGGACGAGGCTGATTATTAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 354
 DB 1141 TGAGGACGAGGCTGATTATTAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082
 QY 355 CTTCGGAAGTGGGACCAAGTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414
 DB 1081 CTTCGGAAGTGGGACCAAGTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
 QY 415 TCTGTTCCGCGCTCTCTCTGAGAGCTTCAAGCCCAAGGCTGCTGCTGCTGCTGCTGCTGCT 474
 DB 1021 TCTGTTCCGCGCTCTCTCTGAGAGCTTCAAGCCCAAGGCTGCTGCTGCTGCTGCTGCTGCT 962
 QY 475 CAGTGAATTTTACCCGGGAGCTGTGACAGTGGCTGGAAGGCAATGAGGCTGCTGCTGCTGCT 534
 DB 961 CAGTGAATTTTACCCGGGAGCTGTGACAGTGGCTGGAAGGCAATGAGGCTGCTGCTGCTGCT 902
 QY 535 GGCCTGGGCTCTGCTATTCTTCCATCCCTCTCCTCAGGAGGAGGCTGGGCGCCAGTGC 594
 DB 901 GGCCTGGGCTCTGCTATTCTTCCATCCCTCTCCTCAGGAGGAGGCTGGGCGCCAGTGC 842
 QY 595 CTACTGAGCTTGAAGGCTGAGGCTGAGTGAAGTCCCAAGAGCTTACAGCTGCTGCTGCTGCT 654
 DB 841 CTACTGAGCTTGAAGGCTGAGGCTGAGTGAAGTCCCAAGAGCTTACAGCTGCTGCTGCTGCT 782
 QY 655 GCATGAAGGAGACACCTGTGATGAAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 691
 DB 781 GCATGAAGGAGACACCTGTGATGAAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 746

RESULT 8

AA83485 standard; cDNA; 863 BP.

AA83485;

13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #19289.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW Food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

PN MO200175067-A2.

XX 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSB-) HYSBQ INC.

PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR P-PDB; ABG19298.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 1, SEQ ID NO 19289; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostic as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AA864197-AA894564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences

XX Sequence 863 BP; 201 A; 277 C; 210 G; 175 T; 0 U; 0 Other;

Query Match 40.8%; Score 595.8; DB 5; Length 863;

Best Local Similarity 94.0%; Pred. No. 2.5e-130;
 Matches 653; Conservative 0; Mismatches 37; Indels 5; Gaps 3;

QY 1 GGCCTGGGCTCTGCTATTCTTCCATCCCTCTCCTCAGGAGGAGGCTGGGCGCCAGTGC 60
 DB 3 GGCCTGGGCTCTGCTATTCTTCCATCCCTCTCCTCAGGAGGAGGCTGGGCGCCAGTGC 62
 QY 61 TGGCCCTGACTAGTGTGCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 DB 63 TGGCCCTGACTAGTGTGCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 122
 QY 121 CACTGGAAACGACATCGTGGTGGTATTATGATGCTCTCTGCTGCTGCTGCTGCTGCTGCTG 180
 DB 123 CACTGGAAACGACATCGTGGTGGTATTATGATGCTCTCTGCTGCTGCTGCTGCTGCTGCTG 182
 QY 181 AGGCAAAAGCCCCCAACTCATCTATTATGAGTCAATGATGAGGCTGCTGCTGCTGCTGCTGCT 240
 DB 183 AGGCAAAAGCCCCCAACTCATCTATTATGAGTCAATGAGGCTGCTGCTGCTGCTGCTGCTGCT 242

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JE;

WPI, 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

Claim 1; Page 3949; 11750bp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

Sequence 1636 BP; 345 A; 470 C; 479 G; 342 T; 0 U; 0 Other;

Query Match 40.0%; Score 584.6; DB 5; Length 1636;

Best Local Similarity 92.6%; Pred. No. 1.2e-127; Mismatches 44; Indels 7; Gaps 2;

Matches 638; Conservative 0; Mismatches 44; Indels 7; Gaps 2;

9 CTCTGCTATTCCTCAACCTCCCTCACTCAGGACAGAGGCTCGGCGCCAGTCTGCCCTGA 68

1552 CACTCTCTTCTCTCACTCTCTCTCACTCAGGACAGAGGCTCGGCGCCAGTCTGCCCTGA 1493

69 CTCAGTCTGCTCTCGTCTGCTCTCTCTGACAGTGCATCACTCTCTGCACTGGAA 128

1492 CTCACCTGCT 1433

129 CCAGAGTCAAGTCTGCT 188

1432 CCGAGAGTCTGCT 1373

189 CCCCCAACTATCTATTAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 248

1372 CCCCCAACTCTGATTAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1313

249 CTGGCTCAAGTCTGCT 308

1312 CTGGCTCAAGTCTGCT 1253

309 AGGCTGATTAATCT 362

1252 AGGCTGATTAATCT 1193

363 CTGGACCAAGTCTGCT 422

1192 CTGGACCAAGTCTGCT 1133

423 CGCCT 482

1132 CGCCT 1073

483 TTTACCCGGAGCTGTGACAGTGTGCTGGAAGGCAAGTGGCAAGTGGCAAGTGGCAAG 542

1072 TTTACCCGGAGCTGTGACAGTGTGCTGGAAGGCAAGTGGCAAGTGGCAAGTGGCAAG 1013

543 TGGAGACCAAGTCTGCT 602

1012 TGGAGACCAAGTCTGCT 953

603 GCTTACCGCCGAGAGTGGAGTCCCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 662

DB 952 GCCTGACCGCTGAGAGTGAAGTCCACAGAGTCAAGTCAAGTCAAGTCAAGTCAAG 893

QY 663 GGAGCACCGTGTGATTAAGCACTGTCTCC 691

DB 892 GGAGCACCGTGTGATTAAGCACTGTCTCC 865

RESULT 12

ABV28405/c

ID ABV28405 standard; CDNA; 1636 BP.

ABV28405;

16-SRP-2002 (first entry)

Human prostate expression marker CDNA 28396.

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; pharmacogenomic marker; gene; ss.

Homo sapiens.

WO200160860-A2.

23-AUG-2001.

20-FEB-2001; 2001WO-US005171.

17-FEB-2000; 2000US-0183319P.

16-MAR-2000; 2000US-0183862P.

25-MAY-2000; 2000US-0207454P.

09-JUN-2000; 2000US-0211314P.

18-JUL-2000; 2000US-0219007P.

13-DEC-2000; 2000US-0255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JE;

WPI, 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

Claim 1; Page 5921; 11750bp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

Sequence 1636 BP; 345 A; 470 C; 479 G; 342 T; 0 U; 0 Other;

Query Match 40.0%; Score 584.6; DB 5; Length 1636;

Best Local Similarity 92.6%; Pred. No. 1.2e-127; Mismatches 44; Indels 7; Gaps 2;

Matches 638; Conservative 0; Mismatches 44; Indels 7; Gaps 2;

9 CTCTGCTATTCCTCAACCTCCCTCACTCAGGACAGAGGCTCGGCGCCAGTCTGCCCTGA 68

1552 CACTCTCTTCTCTCACT 1493

69 CTCAGTCTGCTCTCGTCTGCT 128

1492 CTCACCTGCT 1433


```

QY 129 CCAGCAGTCAGTGGTGGTGTATATATATCTCTCTGTAACCAACAGACCCAGGCAAG 188
DB 1432 CCAGGAGTCTGTGGGATATATACTTGTCTCTGTACACACACACCCAGGCAAG 1373
QY 189 CCCCCAACTCATATTTATGAGGTCAAGTATCGGCTCAGGGGTTCTATCGCTTC 248
DB 1372 CCCCCAACTCATATTTATGAGGTCAAGTATCGGCTCAGGGGTTCTATCGCTTC 1313
QY 249 CTGGCTCAAGTCTGGCAACAGGCTCTCCCTGACCAATCTCTGGGCTCCAGGCTGAGGACG 308
DB 1312 CTGGCTCAAGTCTGGCAACAGGCTCTCCCTGACCAATCTCTGGGCTCCAGGCTGAGGACG 1253
QY 309 AGGCTATTATTAATCTGCTCATATATACAGAAATATCTCT-----CATGCTTGGGA 362
DB 1252 AGGCTATTATTAATCTGCTCATATATACAGAAATATCTCTGGGCTCTGAGGCTGAGGACG 1193
QY 363 CTGGGACCAAGTCAACGCTCTAGTCAAGCCCAAGGCAACCCCACTGTCACTCTGTTCC 422
DB 1192 CTGGGACCAAGTCAACGCTCTAGTCAAGCCCAAGGCTCTGAGGCTGAGGACG 1133
QY 423 CGCCCTCTCTGAGAGGTCCAAAGCAACAGGCTCACTAGTGTCTGATCAGTACT 482
DB 1132 CGCCCTCTCTGAGAGGTCCAAAGCAACAGGCTCACTAGTGTCTGATCAGTACT 1073
QY 483 TCTACCCGAGCTGTGACAGTGGCTGGAAGCAGATGGCAGCCCGTCAAGGCGGAG 542
DB 1072 TCTACCCGAGCGGTGACAGTGGCTGGAAGCAGATGGCAGCCCGTCAAGGCGGAG 1013
QY 543 TGGAGACCAACCAACCTCTCCAAACAGACCAACCAAGTACGCGGCAAGCACTACTGA 602
DB 1012 TGGAGACCAACCAACCTCTCCAAACAGACCAACCAAGTACGCGGCAAGCACTACTGA 953
QY 603 GCTGAGCGCCGAGAGTGAAGTCCCAAGAGCTACGTGCAAGTCAAGCAATGAAG 662
DB 952 GCTGAGCGCCGAGAGTGAAGTCCCAAGAGCTACGTGCAAGTCAAGCAATGAAG 893
QY 663 GGAGCAGCGTGAATGAAGACAGTGTCC 691
DB 892 GGAGCAGCGTGAATGAAGACAGTGTCC 865

RESULT 13
AADS9189/c
ID AADS9189 standard; cDNA; 2667 BP.
XX
AC AADS9189;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human rearranged Ig lambda chain cDNA.
XX
KW Colon cancer; gene therapy; human; Ig; immunoglobulin; ss.
XX
OS Homo sapiens.
XX
PN US2003073105-A1.
XX
PD 17-APR-2003.
XX
PE 29-MAY-2002; 2002US-00158646.
XX
PR 31-MAY-2001; 2001US-0295239P.
XX
PA (LASEK/) LASEK A K W.
XX
PA (SORN/) SORNASSE T.
XX
PI Lasek AKW, Sornasse T;
XX
DR WPI; 2003-605964/57.
XX
PT Novel combination of cDNAs which are differentially expressed in colon
cancer, useful for detecting differential expression of one or more cDNAs

```

```

PT in a sample containing nucleic acid samples.
XX
XX Claim 1, Page 84-85; 88pp; English.
XX
XX The present invention relates to combination of cDNAs which are
CC differentially expressed in colon cancer. The invention is useful for
CC producing and purifying antibody, utilized as markers for treatment
CC efficacy against colon cancer. The invention is also useful for gene
CC therapy. The present sequence is human rearranged Ig lambda chain cDNA
XX
SQ Sequence 2667 BP; 602 A; 804 C; 721 G; 540 T; 0 U; 0 Other;
Query Match 39.8%; Score 581.2; DB 9; Length 2667;
Best Local Similarity 90.2%; Pred. No. 8.8e-127;
Matches 644; Conservative 0; Mismatches 66; Indels 2; Gaps 2;
QY 1 GGCTGGGCTCTGCTATTCCTCACCTCTCTCACTGAGGACAGAGGCTCTGGGCCCAATC 60
DB 2633 GGCTGGGCTCTGCTATTCCTCACCTCTCTCACTGAGGACAGAGGCTCTGGGCCCAATC 2574
QY 61 TGGCTGACTGACTGCTGCTCGGTGTGGGTCTCTGACAGTGCATCATCTCTG 120
DB 2573 TGGCTGACTGACTGCTGCTCGGTGTGGGTCTCTGACAGTGCATCATCTCTG 2514
QY 121 CACTGGAACCAAGCACTGAGTGTGTTATTAATATCTCTCTGTAACCAACAGCACC 180
DB 2513 CACTGGAACCAAGCACTGAGTGTGTTATTAATATCTCTCTGTAACCAACAGCACC 2454
QY 181 AGGCAAGGCCCCCAATCATATTATGAGGTCACTAATGGCCCTCAAGGGTTCTTA 240
DB 2453 AGGCAAGGCCCCCAATCATATTATGAGGTCACTAATGGCCCTCAAGGGTTCTTA 2394
QY 241 TGGCTTCTGCTGCTCAAGTCTGGCAACAGGCTCTCTGACATCTCT-GGGCTCCAGG 299
DB 2393 TGGCTTCTGCTGCTCAAGTCTGGCAACAGGCTCTCTGACATCTCTGGGCTCCAGG 2334
QY 300 CTGAGAGCAGAGCTGATTTACTGTGCTCATATACAGAAAGTACTTCTCATGTCTTG 359
DB 2333 CTGAGAGTGAAGCTGATTTATTTATGATTTGATTTGCAAGAAATGATTTATGCTATTG 2274
QY 360 GAATGGGACCAAGGTCAACCTCTGAGTCAAGCCCAAGCCCAAGCTGCACTGCT 419
DB 2273 GCGAGGAGCCAAAGTTGACCGTCTGAGTCAAGCCCAAGCCCAAGCTGCTGCTGCT 2214
QY 420 TCCGCGCTCTCTGAGAGCTTCAAGCCCAAGCCCAAGCTGCTGATCACTG 479
DB 2213 TCCGCGCTCTCTGAGAGCTTCAAGCCCAAGCCCAAGCTGCTGATCACTG 2154
QY 480 ACTTTTACCCGGAGCTGTGACAGTGTGCTGGAAGGCAATGCGACGCTCAAGGCG 539
DB 2153 ACTTTTACCCGGAGCTGTGACAGTGTGCTGGAAGGCAATGCGACGCTCAAGGCG 2094
QY 540 GAGTGAAGACCAACCAACCTCTCCAAACAGAGCAACCAAGTACGCGGCCAGAGTACC 599
DB 2093 GAGTGAAGACCAACCAACCTCTCCAAACAGAGCAACCAAGTACGCGGCCAGAGTACC 2034
QY 600 TGAAGCTGACGCGCCGAGAGTGAAGTCCCAAGAGTACAGTGTCAAGGCTCAAGCATG 659
DB 2033 TGAAGCTGACGCGCCGAGAGTGAAGTCCCAAGAGTACAGTGTCAAGGCTCAAGCATG 1974
QY 660 AAGGAGAGCAGTGTGATGAAGAGTGTCCCTTACAGAAATGTTCTATAGTT 713
DB 1973 AAGGAGAGCAGTGTGATGAAGAGTGTCCCTTACAGAAATGTTCTATAGTT 1921

RESULT 14
AAS77073
ID AAS77073 standard; cDNA; 889 BP.
XX
AC AAS77073;
XX
DT 13-FEB-2002 (first entry)
XX

```


CC therapy. The present sequence is human Ig (hybridoma H210) anti-hepatitis
CC A Ig lambda chain cDNA

xx Sequence 883 BP, 195 A, 283 C, 220 G, 185 T, 0 U, 0 Other;

Query Match 39.4%; Score 575.2; DB 9; Length 883;
Best Local Similarity 87.9%; Pred. No. 1.8e-125;

Matches 675; Conservative 0; Mismatches 83; Indels 10; Gaps 4;

```
QY 1 GGCCTGGGCTCTGCTATTCTCTACCCCTCTCTACTCAGGCGACAGGGTCTGGGCCAGTGC 60
DB 41 GGCCTGGGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 100
QY 61 TGCCTGACTCAGTCTGCTCCGCTGCTGGGCTCTCTGACAGTCAGTCACTCTCTG 120
DB 101 TGCCTGACTCAGTCTGCTCCGCTGCTGGGCTCTCTGACAGTCAGTCACTCTCTG 160
QY 121 CACTGGAACGACGATCAGTGTGGTTATTAATACTATGTCCTGTGTACCAAGACACC 180
DB 161 CACTGGAAGTGGCGGTGACGTGGTGTATTAATTTGTCCTGTGTATAACACACACC 220
QY 181 AGGCAAGCCCCCAACTCATATTATTAAGTCAATATCGGCCCTCAGGGGTTCTAA 240
DB 221 AGGCAAGCCCCCAACTCATATTATTAAGTCAATATCGGCCCTCAGGGGTTCTAA 280
QY 241 TCGCTTCTGTGCTCAAGTCTGGCAACGCGCTCCCTGACCATCTGTGGCTCCAGGC 300
DB 281 TCGCTTCTGTGCTCAAGTCTGGCAACGCGCTCCCTGACCATCTGTGGCTCCAGGC 340
QY 301 TGAAGACGAGGCTATTTACTGCTCTCATATACA--AGAAGTACTTCTCATGTCTT 357
DB 341 TGAAGATGAGGCTATTTACTGCTCTCATATTTGTAAGTACTTCTGTGTATT 400
QY 358 CGGAAGTGGGACGAGGTCACCGTCTTAGTCAAGCCCAAGGCCCACTAGTGTCTGATCA 417
DB 401 CGGCGAGGAGGACGAGGTCACCGTCTTAGTCAAGCCCAAGGCCCACTAGTGTCTGATCA 460
QY 418 GTTCCCGCCCTCTCTGAGAGTCTCAAGCCCAAGGCCCACTAGTGTCTGATCA 477
DB 461 GTTCCCGCCCTCTCTGAGAGTCTCAAGCCCAAGGCCCACTAGTGTCTGATCA 520
QY 478 TGAAGTCTACCGGAGAGTGTGACAGTGGGCTGGAAGGAGGAGGCGCCGTCAGAGC 537
DB 521 TGAAGTCTACCGGAGAGTGTGACAGTGGGCTGGAAGGAGGAGGCGCCGTCAGAGC 580
QY 538 GGAAGTGAAGACCAACCAACCCCTCCAAACAGAGCAACACAAGTACCGGCGCAGAGCTA 597
DB 581 GGAAGTGAAGACCAACCAACCCCTCCAAACAGAGCAACACAAGTACCGGCGCAGAGCTA 640
QY 598 CCTGAGCTTGAAGCCCGAGAGTGAAGTCCCAAGAACTACAGTCTGCGAGGTCAAGCA 657
DB 641 TCTGAGCTTGAAGCCCGAGAGTGAAGTCCCAAGAACTACAGTCTGCGAGGTCAAGCA 700
QY 658 TGAAGGAGCACGCTGAGTGAAGACAGTGGTCCCTACAGAAATGTTCTATAGTTCTCT 717
DB 701 TGAAGGAGCACGCTGAGTGAAGACAGTGGTCCCTACAGAAATGTTCTATAGTTCTCT 755
QY 718 AACACGCTCAGCCCTCACTACCAATGCGAGACTAGCAAGCTGCAAGG 765
DB 756 --AACCTCAACCCCAACGAGGAGAGTGAAGTGAAGATCCCAAG 801
```

Search completed: April 26, 2004, 16:12:07
Job time : 560.73 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: April 26, 2004, 18:46:58 ; Search time 608.283 Seconds

(Without alignments)
10821.600 Million cell updates/sec

Title: US-10-076-747-8

Sequence: 1 ggcctggcctgtatcc.....ggcagcagcagcagcagc 1460

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 2907579 seqs, 225413464 residues

Total number of hits satisfying chosen parameters: 5815158

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCR_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCRUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*

14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq2:*

17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1459	99.9	1460	US-10-076-747-8
2	809.2	55.5	957	US-10-076-747-7
3	599	41.0	14597	US-10-198-846-13540
4	597.4	40.9	1480	US-09-981-353-146
5	588.6	40.3	1640	US-10-198-846-13206
6	581.2	39.8	2667	US-10-158-646-76
7	575.6	39.4	1597	US-10-198-846-13540
8	575.2	39.4	883	US-10-158-646-73
9	565.4	38.7	729	US-10-461-148-12
10	529.4	36.3	726	US-10-225-108A-13
11	511.8	35.1	2112	US-10-001-857-108
12	505.8	34.6	2635	US-10-198-846-13529
13	498.2	34.1	819	US-10-342-887-1913
14	498.2	34.1	819	US-10-172-118-1913

15	497.2	34.1	605	13	US-09-878-134-264	Sequence 264, App
16	488.6 <td>33.5 <td>9472</td> <td>13 <td>US-09-837-306-194</td> <td>Sequence 194, App</td> </td></td>	33.5 <td>9472</td> <td>13 <td>US-09-837-306-194</td> <td>Sequence 194, App</td> </td>	9472	13 <td>US-09-837-306-194</td> <td>Sequence 194, App</td>	US-09-837-306-194	Sequence 194, App
17	488.6 <td>33.5 <td>9532</td> <td>16 <td>US-10-045-674-451</td> <td>Sequence 451, App</td> </td></td>	33.5 <td>9532</td> <td>16 <td>US-10-045-674-451</td> <td>Sequence 451, App</td> </td>	9532	16 <td>US-10-045-674-451</td> <td>Sequence 451, App</td>	US-10-045-674-451	Sequence 451, App
18	480.6 <td>32.9 <td>543</td> <td>9 <td>US-09-736-457-970</td> <td>Sequence 970, App</td> </td></td>	32.9 <td>543</td> <td>9 <td>US-09-736-457-970</td> <td>Sequence 970, App</td> </td>	543	9 <td>US-09-736-457-970</td> <td>Sequence 970, App</td>	US-09-736-457-970	Sequence 970, App
19	480.6 <td>32.9 <td>543</td> <td>9 <td>US-09-902-941-970</td> <td>Sequence 970, App</td> </td></td>	32.9 <td>543</td> <td>9 <td>US-09-902-941-970</td> <td>Sequence 970, App</td> </td>	543	9 <td>US-09-902-941-970</td> <td>Sequence 970, App</td>	US-09-902-941-970	Sequence 970, App
20	480.6 <td>32.9 <td>543</td> <td>9 <td>US-09-849-626-970</td> <td>Sequence 970, App</td> </td></td>	32.9 <td>543</td> <td>9 <td>US-09-849-626-970</td> <td>Sequence 970, App</td> </td>	543	9 <td>US-09-849-626-970</td> <td>Sequence 970, App</td>	US-09-849-626-970	Sequence 970, App
21	480.6 <td>32.9 <td>543</td> <td>13 <td>US-10-283-017-970</td> <td>Sequence 970, App</td> </td></td>	32.9 <td>543</td> <td>13 <td>US-10-283-017-970</td> <td>Sequence 970, App</td> </td>	543	13 <td>US-10-283-017-970</td> <td>Sequence 970, App</td>	US-10-283-017-970	Sequence 970, App
22	480.6 <td>32.9 <td>543</td> <td>13 <td>US-10-017-754-970</td> <td>Sequence 970, App</td> </td></td>	32.9 <td>543</td> <td>13 <td>US-10-017-754-970</td> <td>Sequence 970, App</td> </td>	543	13 <td>US-10-017-754-970</td> <td>Sequence 970, App</td>	US-10-017-754-970	Sequence 970, App
23	480.6 <td>32.9 <td>543</td> <td>15 <td>US-10-113-872-970</td> <td>Sequence 970, App</td> </td></td>	32.9 <td>543</td> <td>15 <td>US-10-113-872-970</td> <td>Sequence 970, App</td> </td>	543	15 <td>US-10-113-872-970</td> <td>Sequence 970, App</td>	US-10-113-872-970	Sequence 970, App
24	477.2	32.7	1640	16 <td>US-10-198-846-13206</td> <td>Sequence 13206, A</td>	US-10-198-846-13206	Sequence 13206, A
25	473.4	32.4	930	16 <td>US-10-291-265-107</td> <td>Sequence 107, App</td>	US-10-291-265-107	Sequence 107, App
26	473	32.4	848	15 <td>US-10-158-646-70</td> <td>Sequence 788, App</td>	US-10-158-646-70	Sequence 788, App
27	468.2	32.1	915	9 <td>US-09-954-456-788</td> <td>Sequence 788, App</td>	US-09-954-456-788	Sequence 788, App
28	468.2	32.1	915	10 <td>US-09-880-107-3743</td> <td>Sequence 3743, App</td>	US-09-880-107-3743	Sequence 3743, App
29	468.2	32.1	915	10 <td>US-09-960-706-1069</td> <td>Sequence 1069, App</td>	US-09-960-706-1069	Sequence 1069, App
30	468.2	32.1	915	10 <td>US-09-873-319-710</td> <td>Sequence 710, App</td>	US-09-873-319-710	Sequence 710, App
31	468.2	32.1	915	13 <td>US-10-342-887-1912</td> <td>Sequence 1912, App</td>	US-10-342-887-1912	Sequence 1912, App
32	468.2	32.1	915	13 <td>US-10-172-118-1912</td> <td>Sequence 1912, App</td>	US-10-172-118-1912	Sequence 1912, App
33	468.2	32.1	964	13 <td>US-10-363-616-87</td> <td>Sequence 87, App</td>	US-10-363-616-87	Sequence 87, App
34	462.8	31.7	888	15 <td>US-10-158-646-71</td> <td>Sequence 71, App</td>	US-10-158-646-71	Sequence 71, App
35	460	31.5	1590	15 <td>US-10-198-846-12799</td> <td>Sequence 12799, A</td>	US-10-198-846-12799	Sequence 12799, A
36	458	31.4	768	9 <td>US-09-747-669-4</td> <td>Sequence 4, App</td>	US-09-747-669-4	Sequence 4, App
37	458	31.4	768	9 <td>US-09-747-669-5</td> <td>Sequence 5, App</td>	US-09-747-669-5	Sequence 5, App
38	458	31.4	768	15 <td>US-10-290-703-4</td> <td>Sequence 4, App</td>	US-10-290-703-4	Sequence 4, App
39	458	31.4	768	15 <td>US-10-290-703-5</td> <td>Sequence 5, App</td>	US-10-290-703-5	Sequence 5, App
40	454.8	31.2	711	9 <td>US-09-948-4298-9</td> <td>Sequence 9, App</td>	US-09-948-4298-9	Sequence 9, App
41	454.8	31.2	711	13 <td>US-09-758-173-9</td> <td>Sequence 9, App</td>	US-09-758-173-9	Sequence 9, App
42	454.8	31.2	711	14 <td>US-10-124-905-9</td> <td>Sequence 9, App</td>	US-10-124-905-9	Sequence 9, App
43	454.8	31.2	711	14 <td>US-10-073-138-5</td> <td>Sequence 5, App</td>	US-10-073-138-5	Sequence 5, App
44	454.8	31.2	711	15 <td>US-10-124-807-9</td> <td>Sequence 9, App</td>	US-10-124-807-9	Sequence 9, App
45	454.8	31.2	711	15 <td>US-10-291-532-9</td> <td>Sequence 9, App</td>	US-10-291-532-9	Sequence 9, App

ALIGNMENTS

RESULT 1

US-10-076-747-8

Sequence 8, Application US/10076747

Publication No. US20030180726A1

GENERAL INFORMATION:

APPLICANT: Salceda, Susana

APPLICANT: Macina, Roberto

APPLICANT: Hu, Ping

APPLICANT: Recipon, Hervé

APPLICANT: Karra, Kalpana

APPLICANT: Caffrey, Robert

APPLICANT: Sun, Yongming

APPLICANT: Liu, Chenghua

TITLE OF INVENTION: Compositions and Methods Relating to Ovarian Specific Genes and

FILE REFERENCE: DEX-0315

CURRENT APPLICATION NUMBER: US/10/076,747

PRIOR FILING DATE: 2002-02-13

PRIOR APPLICATION NUMBER: 60/268,290

PRIOR FILING DATE: 2001-02-13

PRIOR APPLICATION NUMBER: 60/268,834

NUMBER OF SEQ ID NOS: 129

SOFTWARE: PatentIn version 3.1

SEQ ID NO 8

LENGTH: 1460

TYPE: DNA

ORGANISM: Homo sapien

FEATURE:

NAME/KEY: misc feature

LOCATION: (1022)..(1022)

OTHER INFORMATION: a, c, g or t

US-10-076-747-8

Query Match 99.9%; Score 1459; DB 15; Length 1460;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1460; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTGAGGCTGTATTCTCACCCCTCTCACTCAGGGGACAGGGGCTCTGGGCGCAGTC 60
 Db 1 GGCCTGAGGCTGTATTCTCACCCCTCTCACTCAGGGGACAGGGGCTCTGGGCGCAGTC 60
 QY 61 TCCCTGACTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 Db 61 TCCCTGACTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 QY 121 CACTGGAACAGAGCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Db 121 CACTGGAACAGAGCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 QY 181 AGGCAAAAGCCCCCAAACTCATATTATAGAGTCAAGTCAAGGCTCAGGGGCTTCTTAA 240
 Db 181 AGGCAAAAGCCCCCAAACTCATATTATAGAGTCAAGTCAAGGCTCAGGGGCTTCTTAA 240
 QY 241 TGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 Db 241 TGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 QY 301 TGAGGACGAGGCTGATTTATTACTGCTGCTCATATACAAAGATCTTCTCATGCTTGG 360
 Db 301 TGAGGACGAGGCTGATTTATTACTGCTGCTCATATACAAAGATCTTCTCATGCTTGG 360
 QY 361 AACTGGGACCAAGTCAACCGCTCTAGTCAAGCCCAAGGCCAACCCCACTGCTCACTGCTT 420
 Db 361 AACTGGGACCAAGTCAACCGCTCTAGTCAAGCCCAAGGCCAACCCCACTGCTCACTGCTT 420
 QY 421 CCGGCTCTCTGAGAGCTCCAAAGCCCAAGGCCCACTAGTGTCTGATCAAGTGA 480
 Db 421 CCGGCTCTCTGAGAGCTCCAAAGCCCAAGGCCCACTAGTGTCTGATCAAGTGA 480
 QY 481 CTTCTACCGGAGCTGTGAGAGCTCCAAAGCCCAAGGCCCACTAGTGTCTGATCAAGTGA 540
 Db 481 CTTCTACCGGAGCTGTGAGAGCTCCAAAGCCCAAGGCCCACTAGTGTCTGATCAAGTGA 540
 QY 541 AGTGAAGACCAACCAACCTCCAAAGAGCAACCAAGTCAAGGCCCAAGGCCCACTAGTGT 600
 Db 541 AGTGAAGACCAACCAACCTCCAAAGAGCAACCAAGTCAAGGCCCAAGGCCCACTAGTGT 600
 QY 601 GAGCTGAGGCGCCGAGAGTGAAGTCCCAAGAGCTCAAGCTCAGGATCAGCATGA 660
 Db 601 GAGCTGAGGCGCCGAGAGTGAAGTCCCAAGAGCTCAAGCTCAGGATCAGCATGA 660
 QY 661 AGGAGACACCGTGAATGAAGACAGTGTCCCTACAGAAATGTTCTATAGGTTCTTAA 720
 Db 661 AGGAGACACCGTGAATGAAGACAGTGTCCCTACAGAAATGTTCTATAGGTTCTTAA 720
 QY 721 AGGCTACGCGCCCACTACCAATGAGGAGACTAGCAGCTGCAAGGATCCCAAGGAGAGG 780
 Db 721 AGGCTACGCGCCCACTACCAATGAGGAGACTAGCAGCTGCAAGGATCCCAAGGAGAGG 780
 QY 781 GGTCTCTCATCCACACACACAGAGGCGAGTCAAGGCTTATCATGCGCATGTGAC 840
 Db 781 GGTCTCTCATCCACACACACAGAGGCGAGTCAAGGCTTATCATGCGCATGTGAC 840
 QY 841 GTCTATTAAGAGCGCTTACAAATTAAGATTTTGTGCACTTGGTGAATGTCTTACATCA 900
 Db 841 GTCTATTAAGAGCGCTTACAAATTAAGATTTTGTGCACTTGGTGAATGTCTTACATCA 900
 QY 901 TAAACAAAGACATCTCACTACACAGGAGTCAAGGCTTATCATGCGCATGTGAC 960
 Db 901 TAAACAAAGACATCTCACTACACAGGAGTCAAGGCTTATCATGCGCATGTGAC 960
 QY 961 GGCACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 Db 961 GGCACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 QY 1021 AACTATGAGTGAATTTGGCCGCGGAGTACAGACATGATTTTCCCGCGGCTGATCG 1080
 Db 1021 AACTATGAGTGAATTTGGCCGCGGAGTACAGACATGATTTTCCCGCGGCTGATCG 1080
 QY 1081 AATATGCGGTACGCGGACCAAAATTCGCCCGGAGCTACAGACCTTAAAAAATGTGACGCG 1140

Db 1081 AATATGCGGTACGCGGACCAAAATTCGCCCGGAGCTACAGACCTTAAAAAATGTGACGCG 1140
 QY 1141 GCAGCGCGCGGAGAGAGAGGCTTTGAGAGGATGCCAAACACAGAGGCGCGCAGACCGGCA 1200
 Db 1141 GCAGCGCGCGGAGAGAGAGGCTTTGAGAGGATGCCAAACACAGAGGCGCGCAGACCGGCA 1200
 QY 1201 GGCAGCCCGGCGGAGAGAGGCGCCCACTGCGGAGAGGCGAGGAGAGAGAGATACAGAGA 1260
 Db 1201 GGCAGCCCGGCGGAGAGAGGCGCCCACTGCGGAGAGGCGAGGAGAGAGATACAGAGA 1260
 QY 1261 CCGAGCCCAACCGCGCGCGCTTGCATGAGACGAGAGCGGCGAGGAGAGCGCAGACCGGAA 1320
 Db 1261 CCGAGCCCAACCGCGCGCGCTTGCATGAGACGAGAGCGGCGAGGAGAGCGCAGACCGGAA 1320
 QY 1321 GCAGCGCGCGGAGAGAGAGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
 Db 1321 GCAGCGCGCGGAGAGAGAGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
 QY 1381 CAGCGCGAG 1440
 Db 1381 CAGCGCGAG 1440
 QY 1441 GGCAGCAGAGCCACCGAGCT 1460
 Db 1441 GGCAGCAGAGCCACCGAGCT 1460

RESULT 2

US-10-076-747-7

Sequence 7, Application US/10076747

Publication No. US20030180726A1

GENERAL INFORMATION:

APPLICANT: Salceda, Susana

APPLICANT: Macina, Roberto

APPLICANT: Hu, Ping

APPLICANT: Recipon, Hervé

APPLICANT: Kariya, Kalpana

APPLICANT: Cafferykey, Robert

APPLICANT: Sun, Yongming

APPLICANT: Liu, Chengshua

TITLE OF INVENTION: Compositions and Methods Relating to Ovarian Specific Genes and

FILE REFERENCE: DBX-0315

CURRENT APPLICATION NUMBER: US/10/076,747

CURRENT FILING DATE: 2002-02-13

PRIOR APPLICATION NUMBER: 60/268,290

PRIOR FILING DATE: 2001-02-13

PRIOR APPLICATION NUMBER: 60/268,834

PRIOR FILING DATE: 2001-02-15

NUMBER OF SEQ ID NOS: 129

SOFTWARE: PatentIn version 3.1

SEQ ID NO 7

LENGTH: 957

TYPE: DNA

ORGANISM: Homo sapien

FEATURE:

NAME/KEY: misc.feature

LOCATION: (519)..(519)

OTHER INFORMATION: a, c, g or t

US-10-076-747-7

Query Match 55.5%; Score 809.6; DB 15; Length 957;
 Best Local Similarity 96.5%; Pred. No. 2.2e-230;
 Matches 864; Conservative 0; Mismatches 19; Indels 12; Gaps 3;

QY 578 AAGTACGCGCGCAGAGCTACTGAGCTGAGCGCC---GAGCAGTGAAGTCCGACAG 633
 Db 63 AGGTACGCGCGGTGAGAGCTATCTGAGCTGAGCGCTGATGAGTGAAGTCCGACAG 122
 QY 634 AAGCTACAGTCCAGAGTAC-----GATGAAGAGAG-CACCGTGAATGAAGCAGT 685
 Db 123 AAGCTACAGCTTGACAGAGCTTATGCTGATGAAGAGAGACCGGTGATGAAGCAGT 182

```

QY 686 GGTCCCTACAGAAATGTTCTATAGTTCTCTAACAACGCTCAGCCCCCACTACCAATGCG 745
DB 183 GGTCCCTACAGAAATGTTCTATAGTTCTCTAACAACGCTCAGCCCCCACTACCAATGCG 242
QY 746 GAGACTAGCAAGCTGCAAGGATCCCAAGGAGAGGGGTCTCTCATCCACACCAACAAG 805
DB 243 GAGACTAGCAAGCTGCAAGGATCCCAAGGAGAGGGGTCTCTCATCCACACCAACAAG 302
QY 806 GGGGAGTCAAGCCCTTATCATGCGGCAATGCAACGTCAGTAAAGCCGCTCAAAATA 665
DB 303 GGGGAGTCAAGCCCTTATCATGCGGCAATGCAACGTCAGTAAAGCCGCTCAAAATA 362
QY 866 GATATTCGACCTTGTTGAATGTCTTACATACATAACAACAACATCTCAATCACTACA 925
DB 363 GATATTCGACCTTGTTGAATGTCTTACATACATAACAACAACATCTCAATCACTACA 422
QY 926 CGGAGTGAATGATCAACCGGTCCGTGCGGGCAATGCCACTTGCTCTGCTGCTCAAT 985
DB 423 CGGAGTGAATGATCAACCGGTCCGTGCGGGCAATGCCACTTGCTCTGCTGCTCAAT 482
QY 986 GATAGTCAATAGTACACACCGGTGTGCGGGCAACATGATGATGATTTGCGCGCGG 1045
DB 483 GATAGTCAATAGTACACACCGGTGTGCGGGCAACATGATGATGATTTGCGCGCGG 542
QY 1046 ATCAACAGACCATGATTTCCCGGTGCTGATCGAATATGCGGTAACGCGCATCAAAAT 1105
DB 543 ATCAACAGACCATGATTTCCCGGTGCTGATCGAATATGCGGTAACGCGCATCAAAAT 602
QY 1106 CGCCCGGAGCTACAGACCTTAAAAAGTTGACCGCGGAGCGGCGCAAGAACAGCTTTGG 1165
DB 603 CGCCCGGAGCTACAGACCTTAAAAAGTTGACCGCGGAGCGGCGCAAGAACAGCTTTGG 662
QY 1166 ACGGAAATGCCAAACACAGAGGGCGCGACACCGGACGCGACCCCGGCGAGAGAGCCCA 1225
DB 663 ACGGAAATGCCAAACACAGAGGGCGCGACACCGGACGCGACCCCGGCGAGAGAGCCCA 722
QY 1226 CTGCGGCAAGGCGGCGGAGAGACATACAGAGAGCGGACCAACGCGCGCTTGCA 1285
DB 723 CTGCGGCAAGGCGGCGGAGAGACATACAGAGAGCGGACCAACGCGCGCTTGCA 782
QY 1286 TGAACAGGAGAGCGGCGGAGAGAGCGGACACCGGAGAGAGCGGCGGCGGCA 1345
DB 783 TGAACAGGAGAGCGGCGGAGAGAGCGGACACCGGAGAGAGCGGCGGCGGCA 842
QY 1346 AGCCACACGCGCTTACGCGCTGCGCAACGCGCGGCGCAACGCGCGGCGGAGCA 1405
DB 843 AGCCACACGCGCTTACGCGCTGCGCAACGCGCGGCGCAACGCGCGGCGGAGCA 902
QY 1406 CCGGAGAGAGCGGCGGAGAGAGCGGACCGGCGGCGGCGGAGAGAGCGGAGCACT 1460
DB 903 CCGGAGAGAGCGGCGGAGAGAGCGGACCGGCGGCGGCGGAGAGAGCGGAGCACT 957

```

```

RESULT 3
US-10-198-846-13540/c
; Sequence 13540, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198, 846
; PRIORITY FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306, 220
; PRIORITY FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13540

```

```

; LENGTH: 1597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-13540
Query Match
Best Local Similarity 92.6%; Pred. No. 1,3e-167;
Matches 640; Conservative 0; Mismatches 50; Indels 1; Gaps 1;
41.0%; Score 599; DB 15; Length 1597;
QY 1 GGCTTGAGCTCTGATTTCTTCAACCTTCTCACTCAGGGCAAGGGTCTTGGGCCAATC 60
DB 1528 GGCTTGAGCTCTGATTTCTTCAACCTTCTCACTCAGGGCAAGGGTCTTGGGCCAATC 1469
QY 61 TGCCCTGACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB 1468 TGCCCTGACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1409
QY 121 CACTGGAACAGACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB 1408 CACTGGAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1349
QY 181 AGGCAAAAGCCCAACATCATTTATGAGTGAATGAGGCTTCAATGAGGCTTCAATGAGG 240
DB 1348 AGGCAAAAGCCCAACATCATTTATGATGATGATGATGATGATGATGATGATGATGATG 1289
QY 241 TCGCTTCTGAGCTCAGATCTGAGCAACAGGCTCCTGACATCTCTGAGCTCAGAGC 300
DB 1288 TCGCTTCTGAGCTCAGATCTGAGCAACAGGCTCCTGACATCTCTGAGCTCAGAGC 1229
QY 301 TGAGAGAGAGCTGATTTATCTGCTGCTCATATACAAAGTACTTCTCATGCTTCTG 360
DB 1228 TGAGAGAGAGCTGATTTATCTGCTGCTCATATACAAAGTACTTCTCATGCTTCTG 1169
QY 361 AACTGGAACAGGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 1168 AACTGGAACAGGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1109
QY 421 CCGGCTCTCTTGAAGGCTTCAAGCCCAAGGCGCAACTAGTGTCTGATCAGTGA 480
DB 1108 CCGGCTCTCTTGAAGGCTTCAAGCCCAAGGCGCAACTAGTGTCTGATCAGTGA 1049
QY 481 CTTTACCCGAGAGCTGACAGTGGCTGAGAGGAGATGAGGAGGCGGCGGCAAGGCGG 540
DB 1048 CTTTACCCGAGAGCTGACAGTGGCTGAGAGGAGATGAGGAGGCGGCGGCGGCAAGGCGG 989
QY 541 AGTGAAGACCAACCAACCTTCAACAGAGCAACACAGTACGCGGCGAGAGCTACT 600
DB 988 AGTGAAGACCAACCAACCTTCAACAGAGCAACACAGTACGCGGCGAGAGCTACT 929
QY 601 GAGCTGAGCGGCGGAGGCTGAGTCCCAAGAGCTACAGAGCTACAGGCTCAGGAGTGA 660
DB 928 GAGCTGAGCGGCGGAGGCTGAGTCCCAAGAGCTACAGAGCTACAGGCTCAGGAGTGA 869
QY 661 AGGAGACACGCTGATGAGACAGTGTCCC 691
DB 868 AGGAGACACGCTGATGAGACAGTGTCCC 839

```

```

RESULT 4
US-09-981-353-146/c
; Sequence 146, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Iasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; PRIORITY FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 146
; LENGTH: 1480

```

```

; TYPE:DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 1292606.3
; NAME/KEY: unsure
; LOCATION: 134,199, 206
; OTHER INFORMATION: a, t, c, g, or other
US-09-981-353-146

```

Query Match	40.9%;	Score 597.4;	DB 9;	Length 1480;
Best Local Similarity	93.1%;	Pred. No. 3.9e-167;		
Matches 649;	Conservative	0;	Mismatches 41;	Inbels 7; Gaps 2

OY	66	GGCTCTGGGCTCTGCTATTTCCTCACCCTCTCAGTCAGAGGACAGAGGCTCTGGGCTCCAGTTC	60
Db	1441	GGCTCTGGGCTCTGCTATTTCCTCACCCTCTCAGTCAGAGGACAGAGGCTCTGGGCTCCAGTTC	1382
OY	61	TGGCCGTGACTCACTGCTGCTCCGCTGCTGGGCTCTCTCGAGCAGTTCGATACCACTCTCTCG	120
Db	1381	TGGCCGTGACTCACTGCTGCTCCGCTGCTGGGCTCTCTCGAGCAGTTCGATACCACTCTCTCG	1322
OY	121	CACGTGAAACGAGCAGTCACGTTGGTGGTTAAATCATGTCCTCGTACCAACAGACACC	180
Db	1321	CACGTGAGCCAGAGAGTGCATTGGTGCTTAAATAAATGTCTCTCGTACCAACACACC	1262
OY	181	AGGCAAAAGCCCCCAATCTCATCTATTATGAGTCAGTATATGAGCCCTCAGGGGTTTCTAA	240
Db	1261	AGGCAAAAGCCCCCAATCTTATATTATGCTGTCGTATCGAGTCGGGCCCTCAGGGGTTTCAAC	1202
OY	241	TGCGTCTCTGGGCTCCAGTCGGGAAACAGGGCCCTCCAGATCTCTGGGCTCCAGGC	300
Db	1201	TGCGTCTCTGGGCTCCAGTCGGGAAACAGGGCCCTCCAGATCTCTGGGCTCCAGGC	1142
OY	301	TGAGAACGAGCTGATTATTACTGCTGCTCATATACAAGATGACTTCTCAT-----GT	354
Db	1141	TGAGAACGAGCTGATTATTACTGAGCTCATATATGATTCAGATAGCAATCTGAGGGGTT	1082
OY	355	CTTTCGGAACTGGGACCAAGTCACTGCTCTTAGTTCAGCCCAAGGCCAACCCCACTGTCTAC	414
Db	1081	CTTTCGAACTGGGACCAAGGCTTCGCTCTTAGTTCAGCCCAAGGCCAACCCCACTGTCTAC	1022
OY	415	TCTGTTCCGCGCCTCCTCTGAGGAGCTCCAAGCCAAACAAGCCACACTAGTGTGTAT	474
Db	1021	TCTGTTCCGCGCCTCCTCTGAGGAGCTCCAAGCCAAACAAGCCACACTAGTGTGTAT	962
OY	475	CAGTGACTTCTAACCCGGGAGCTGTGACAGTGGCCTGGAAGGCAAGATGGCAGCCCCCTCAA	534
Db	961	CAGTGACTTCTAACCCGGGAGCTGTGACAGTGGCCTGGAAGGCAAGATGGCAGCCCCCTCAA	902
OY	535	GGGGGGAAGTGGAGACCAACCAACCTCCAAACAAGAGCAACAACAAGTAGGGGCGACGAG	594
Db	901	GGGGGGAAGTGGAGACCAACCAACCTCCAAACAAGAGCAACAACAAGTAGGGGCGACGAG	842
OY	595	CTACCTTGAGCTTGACGCCGCCGAGCACTGGAAAGTCCACAAGAACTTACAGTCCCAAGTCA	654
Db	841	CTACCTTGAGCTTGACGCCGCCGAGCACTGGAAAGTCCACAAGAACTTACAGTCCCAAGTCA	782
OY	655	GCATGAAGGGAGCACTCTGAGTGAAGAACAAGTGTCCC	691
Db	781	GCATGAAGGGAGCACTCTGAGTGAAGAAGAAGTGTCCCC	746

RESULT 5
US-10-198-846-13206/c
; Sequence 13206, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinhmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

```

? TITLE OF INVENTION:  FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
? TITLE OF INVENTION:  THERAPY OF BREAST CANCER
? FILE REFERENCE:  MRI-049
? CURRENT APPLICATION NUMBER:  US/10/198,846
? CURRENT FILING DATE:  2002-07-18
? PRIOR APPLICATION NUMBER:  60/306,220
? PRIOR FILING DATE:  2001-07-18
? NUMBER OF SEQ ID NOS:  14084
? SOFTWARE:  FastSeq for Windows Version 4.0
? SEQ ID NO 13206
? LENGTH:  1640
? TYPE:  DNA
? ORGANISM:  Homo sapiens
US-10-198-846-13206

```

Query Match	40.3%;	Score 588.6;	DB 15;	Length 1640;
Best Local Similarity	89.4%;	Pred. No. 1.7e-164;		
Matches 657;	Conservative 0;	Mismatches 74;	Indels 4;	Gaps 2;

Qy	61	TGCGCTGATCTAGTCTGCCTCCGCTCTCGGGTCTTCCTGGAAAGTCATACACATCTCTG	120
Dp	1507	TGCGCTGATCTAGCTCTCCCTCCTCAGTGTCCGGGTCTTCCTGGAAAGTCATACACATCTCTG	144
Qy	121	CACGTGAACACAGACAGTCAGTGTGGTGTATTAACATATGTCTCTCGTATCAACAGCACCC	180
Dp	1447	CATGTGAACACAGATGAATGTTGGTGTATTAACATGTTGTCTCGTATCAACAGCACCC	138
Qy	181	AGGCCAAGCCCCCAAATCATATATTATTAAGAGTCACTAATGTGCGCCCTCAGGGGTTCTTA	240
Dp	1387	AGGCCAAGCCCCCAAATCTCTATTAAGATGCACTAAGGGCGCCCTCAGGGGTTCTTA	122
Qy	241	TGCGCTCTCTGGCTCAAGTCTGGACAACGAGCCCTCCAGCACTCTCTGG3CTCCAG3C	300
Dp	1327	TGCGCTCTCTGGCTCCAAAGTCTGGACAACGAGCCCTCCAGCACTCTCTGG3CTCCAG3C	126
Qy	301	TGAGACGAGGCTGATTATTACTGCTCTCATATPACAAAGATGTACTTCTCATGTCTTG	360
Dp	1267	TGAGAGTGAAGGCTGATTATTACTGCTCTCATATGAGATATAGCACTTG3GTGTTCG	120
Qy	361	AACGTGAACCAAGTCAACGCTCTGTAGTCAAGCCCAAGGCCAACCCTGATCACTGT	420
Dp	1207	CGAGGGAACAAAGCTGACGCTCTGTAGTCAAGCCCAAGGCTGCCCCCTCGATCACTGT	114
Qy	421	CCGCGCTCTCTTGAGAGGCTCCAAAGCCAAACAAGGCCACACTAGTGTGTCTGATCACTGA	480
Dp	1147	CCGCGCTCTCTTGAGAGGCTTCAAGCCAAACAAGGCCACACTGTGTGTCTCATATGTA	108
Qy	481	CTTCTTACCGGGAAGCTGTGACAGTGGCTTGAAGGAGATGGAGCCCGCCCAAGGGCGG	540
Dp	1087	CTTCTTACCGGGAAGCGGTGACAGTGGCTTGAAGGAGATGACAGCCCGCCCAAGGGCGG	102
Qy	541	AGTGAAGACACCAAACTCTCCAAACAGAGCAACAACAATGCGGGCCAGCACTACT	600
Dp	1027	AGTGAAGACCAACAACCTCTCCAAACAAGCAACAACAATGCGGGCCAGCACTACT	96
Qy	601	GAGCTTGAAGCCCGGAGCAAGTGAATCCCAACAAAGTACAGCTGGCCAGGTCACGGCATGA	660
Dp	967	GAGCTTGAAGCTTGAAGCAAGTGAATCCCAACAAAGTACAGCTGGCCAGGTCACGGCATGA	908
Qy	661	AGGAGACACCGTGTGATGAAGACAGTGTCTCCCTACAGAAATGTTCTATAGATTCTCTAAC	720
Dp	907	AGGAGACACCGTGTGAAGAAACAGTGTCCC---TACAAATGTTCTATAGATTCTCTAAC	852
Qy	721	ACGCTACGCTCCAC 735	
Dp	851	CTCACCCCTCCAC 837	

RESULT 6

US-10-158-646-76/c
; Sequence 76, Application US/10158646
; Publication No. US20030073105A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Sornase, Thierly
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0030-1 US
; CURRENT APPLICATION NUMBER: US/10/158,646
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/295,239
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PERL Program
; SEQ ID NO 76
; LENGTH: 2667
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Inycle ID No. US20030073105A1 1329881.6
US-10-158-646-76

Query Match 39.8%; Score 581.2; DB 15; Length 2667;
Best Local Similarity 90.2%; Pred. No. 3.2e-162;
Matches 644; Conservative 0; Mismatches 68; Indels 2; Gaps 2;

```
QY 1 GGCCTGGGCTGCTGATTCCTCAACCCCTCCCTCACTGAGGAGGACAGGGGCTCGGGGCCAAGTC 60
DB 263 GGCCTGGGCTGCTGATTCCTCAACCCCTCCCTCACTGAGGAGGACAGGGGCTCGGGGCCAAGTC 2574
QY 61 TGCCCTGACTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB 2573 TGCCCTGACTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2514
QY 121 CACTGGAACCGACGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB 2513 CACTGGAACCGACGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2454
QY 181 AGGCAAGCCCCCAACTCATCTATTATGAGTCAATGAGTCAAGGCTCCAGGGGTTCTAA 240
DB 2453 AGGCAAGCCCCCAACTCATCTATTATGAGTCAATGAGTCAAGGCTCCAGGGGTTCTAA 2394
QY 241 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 299
DB 2393 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2334
QY 300 CTGAGGACGAGGCTGATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 359
DB 2333 CTGAGGACGAGGCTGATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2274
QY 360 GAAGTGGAGCCAGGATGACCGTCTAGTCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCA 419
DB 2273 GCGAGGAGCCAGGATGACCGTCTAGTCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCA 2214
QY 420 TCCCGCCCTCTCTGAGGAGTCTCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCA 479
DB 2213 TCCCGCCCTCTCTGAGGAGTCTCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCA 2154
QY 480 ACTTCTACCCGAGGAGTGTGACATGAGTCAAGGAGGAGTCAAGGAGGAGTCAAGGAGGAGTCAAGGAG 539
DB 2153 ACTTCTACCCGAGGAGTGTGACATGAGTCAAGGAGGAGTCAAGGAGGAGTCAAGGAGGAGTCAAGGAG 2094
QY 540 GAGTGGAGACCAACCAACCTCTCAACAGAGCAACCAACCAACCAACCAACCAACCAACCAACCAAC 599
DB 2093 GAGTGGAGACCAACCAACCTCTCAACAGAGCAACCAACCAACCAACCAACCAACCAACCAACCAAC 2034
QY 600 TGAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 659
DB 2033 TGAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1974
QY 660 AAGGAGCAACCGTGAATGAAGACAGTGTCCCTCAAGAAATGTTCTAATAGTT 713
```

DB 1973 AAGGAGCAACCGTGAATGAAGACAGTGTCCCTCAAGAAATGTTCTAATAGTT 1921

RESULT 7
US-10-198-846-13540
; Sequence 13540, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13540
; LENGTH: 1597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-13540

Query Match 39.4%; Score 575.6; DB 15; Length 1597;
Best Local Similarity 88.6%; Pred. No. 1.2e-160;
Matches 684; Conservative 0; Mismatches 74; Indels 14; Gaps 5;

```
QY 1 GGCCTGGGCTGCTGATTCCTCAACCCCTCCCTCACTGAGGAGGACAGGGGCTCGGGGCCAAGTC 60
DB 37 GGCCTGGGCTGCTGATTCCTCAACCCCTCCCTCACTGAGGAGGACAGGGGCTCGGGGCCAAGTC 96
QY 61 TGCCCTGACTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB 97 TGCCCTGACTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 156
QY 121 CACTGGAACCGACGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB 157 CACTGGAACCGACGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 216
QY 181 AGGCAAGCCCCCAACTCATCTATTATGAGTCAATGAGTCAAGGCTCCAGGGGTTCTAA 240
DB 227 AGGCAAGCCCCCAACTCATCTATTATGAGTCAATGAGTCAAGGCTCCAGGGGTTCTAA 276
QY 241 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 277 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 336
QY 301 TGAAGACGAGGCTGATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 354
DB 337 TGAAGACGAGGCTGATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 396
QY 355 CTTGGAACTGGAGCAACGATGACCGTCTAGTCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCA 414
DB 397 ATTGGCGAGGAGGACCAAGCTGACCGTCTAGTCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCA 456
QY 415 TGTGTTCCCGCCCTCTCTGAGGAGTCTCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCA 473
DB 457 TGTGTTCCCGCCCTCTCTGAGGAGTCTCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCA 516
QY 474 TCAAGTACTTCAACCCGAGGAGTGTGACATGAGTCAAGGAGGAGTCAAGGAGGAGTCAAGGAGGAGTCAAGGAG 533
DB 517 TCAAGTACTTCAACCCGAGGAGTGTGACATGAGTCAAGGAGGAGTCAAGGAGGAGTCAAGGAGGAGTCAAGGAG 576
QY 534 AGGCGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 593
DB 577 AGGCGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 636
QY 594 GCTACTGAGGCTGACGCGCCGAGGAGTGGAAAGTCCCAAGAAAGTCAAGCTGACAGGTCA 653
```


Db 318 TGAAGATGAGGCTGATTATTACTGCTGCTCATATGCAAGGCGACTACACCCCGGCGTGGT 3777

Oy 355 CTTCGGAACTGGGAGCCAAAGTCAACGGTCCCTAGGTAGCCCAAGGCCAACCCCATGTGCAC 4149

Db 378 TTTCGGGGGAGGAGCCAAAGTGAACCGTCTTAAGTACGCCAAGGTGCCCCCTGTGTAC 4378

Oy 415 TCTGTTCCGCGCCTCCTCTGAGGAGCTCCAAAGCCAAAGGCCACACTGTGTGTGAT 4741

Db 438 TCTGTTCCCGCCCTCCTCTGTAGGAGCTTCAAAGCAAGGCGACACATGTTGTCTCAT 497

Oy 475 CAGGAGTTTTCACCGGAGCTGTGCACATGGCCCTGGAAGGCAATGGCAACCCCTGCAC 534

Db 498 AAGGATTTTACCCGGAGCGGTGCATGTGCTTGAAGGCAAGTACACCCCGTCAA 5578

Oy 535 GCGCGGAGTGGAGACCAACCAACCTCTCCAAACAGAGCAACAAGTACGCGGCGACAG 594

Db 558 GCGCGGAGTGGAGACCAACCAACCTCTCCAAACAAAGCAAGTACGCGGCGACAG 617

Oy 595 CTACCTTGAAGCTTGAACGCCCGAGCAGTGGAAATCCCAAGAAAGCTACAGCTTGCAGGTAC 654

Db 618 CTACCTTGAAGCTTGAACGCCCTGAGCAGTGGAAAGTCCCAAGAAAGCTACAGCTTGCAGGTAC 6777

Oy 655 GCATGAAGGAGCAACCGTGAATGAACAGAGTGTCCC 691

Db 678 GCATGAAGGAGCAACCGTGA-CAACACAGTGGCCCC 713

```

RESULT 10 -108A-13
; Sequence 13, Application US/10225108A
; Publication No. US20030157112A1
; GENERAL INFORMATION:
; APPLICANT: HOOVER, Craig
; APPLICANT: DIETZSCHOLD, Bernhard
; TITLE OF INVENTION: Recombinant Antibodies, and Compositions
; TITLE OF INVENTION: and Methods for Making Them
; FILE REFERENCE: 8321-110
; CURRENT APPLICATION NUMBER: US/10/025,108A
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: US 09/848,832
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/204,518
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: US 60/314,023
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 726
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-108A-13

```

Query Match	36.3%	Score 529.4	DB 15	Length 726
Best Local Similarity	89.8%	Pred. No. 5.3e-147		
Matches 626	Conservative 0	Mismatches 61	Indels 10	Gaps 5

[illegible]

QY 241 TCGCTTCTCGAGCTCCAAAGTGGCAAACAGGCCCTCCCTGACCAATCTCTGGAGTCCAGAGC 30.00

Db 258 TCGCTTCTCGAGCTCCAAAGTGGCAAACAGGCCCTCCCTGACCAATCTCTGGAGTCCAGAGC 31.7

QY 301 TGAGAGCAGAGCTGATTATTACTGCTGCTCATATACAG-----AGTACTTCTCATGT 35.4

Db 318 TGAGAGTGAAGGCTGATTATTACTGCTGCTCATATACAGGCACTACACCCGAGGCGTGT 37.77

QY 355 CTTCGGAACTGGGAACTCAAGGTCAAGCCTCTAAGTGAAGCCAAAGGCCAAACCCCACTGTAC 41.4

Db 378 TTTCGGGGAAGGGAACCAAGCTCAAGCTCTTAAGTCAAGGCCCAAGGCTGCCCCCTCGATAC 43.7

QY 415 TCTGTTCCGCGCTCTCTGAGAGCTCCAAAGCCAAAGGCCCACTAAGTGTCTGAT 47.4

Db 438 TCTGTTCCGCGCTCTCTGAGAGCTCTCAAGCCAAAGGCCCACTAAGTGTCTCAT 49.7

QY 475 CAGTGACTTCTACCCGGGAGCTGTGTACAGTGGCTCTGGAAGGCAATGGCAAGCCCCCTCAA 53.4

Db 498 AAGTGACTTCTACCCGGGAGCTGTGTACAGTGGCTCTGGAAGGCA-ATAGCAAGCCCCCTCAA 55.6

QY 535 GAGGGAGTGAAGCAACCAACCCCTCCAAAACAGACCAACAAACATGACGCGGCACAG 59.4

Db 557 GGGGGAGTGAAGCAACCAACCAACCCCTCCAAAACAAAGCAACAC-AGTACGGGCGACAG 61.5

QY 595 CTACCTGAAGCTTGAGCGCCCGAGCAATGGAAGTCCCAAGAAAGCTACAGCTGCCAGGTCAAC 65.4

Db 616 CTACCTGAAGCTTGAGCGCCCGAGCAATGGAAGTCCCAAGAAAGC-ACAGCTGCCAGGTCAAC 67.4

QY 655 GCATGAAGGAGCAACCGTGATGAAGAAGTGTCTCC 69.1

Db 675 GCATGAAGGAGCAACCGTGA-AGAAACAAGTGTCTCC 71.0

```

RESULT 11
US-10-001-857-108
; Sequence 108, Application US/10001857
; Publication No. US20020183500A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Proteins
; FILE REFERENCE: DEX-0273
; CURRENT APPLICATION NUMBER: US/10/001,857
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,054
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 108
; LENGTH: 2112
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2005)..(2005)
; OTHER INFORMATION: a, c, g or t
US-10-001-857-108

```

Query Match	35.1%	Score 511.8;	DB 14;	Length 2112;
Best Local Similarity	87.2%;	Pred. No. 1.3e-141;		
Matches 585;	Conservative 0;	Mismatches 82;	Indels 4;	Gaps 2.

Qy	Qy	Qy	Qy
24	CCCTCTCACTCAGGACAGAGGCTCTTGGGCCCAGTCTGGCCGTACCTCTTGGCCG	83	
Db	779	CCCTTGGCCCAACGCTCTCAGGCTCTTGGGCCAGTGTGTGTCGACGACGCGCCCTCAG	538
Qy	84	TGTCCTGGCTCTCTTGGACAGTCGATCAACACTCTCTGCACTGGACACGAGTGCATGTTG	143
Db	839	TGTCCTGGGCCCCCAGGACAGAGGTGACACACTCTCTGCACTGGAGAGGCTCCMAATGCG	898

QY 144 G---TGGTATATACTATGTCTCCGTGGTACCAACAGACCCAGCAAGCCCCCAAACTCA 200
 Db 899 GGGGAGGTTATGACTATGTAACACTGGTACCAGACAGTTCCAGAAACAGCCCCCAACTCA 958
 QY 201 TCAATTATAGAGTCACTAATCGGCCCTCAGGGGTTTCTAATGCTTCTTGCTCCAAAGT 260
 Db 959 TGAATTATAGAGTCACTAATCGGCCCTCAGGGGTTTCTAATGCTTCTTGCTCCAAAGT 1018
 QY 261 CTGGCAACAAGGCTCCCTGACCATCTGGGCTCCAGGCTGAGACAGAGGTGATTTAT 320
 Db 1019 CTGGCAACAAGGCTCCCTGACCATCTGGGCTCCAGGCTGAGACAGAGGTGATTTAT 1078
 QY 321 ACTGCTGCTCATATACAAAGTACTTCTCATGCTCTTGGAACTGGGAACCAAGTCAACG 380
 Db 1079 ACTGCTGCTCATATGAGGACGACTACACTTGGGTGTTGGCGGAGGACCAAGTCAACG 1138
 QY 381 TCTTAGGTAGCCCAAGGCAACCCCACTGTCACTCTGTCCGCCCTCTCTTGAAGAGC 440
 Db 1139 TCTTAGGTAGCCCAAGGCTGCCCCCTCGGTCACTCTGTCCGCCCTCTCTTGAAGAGC 1198
 QY 441 TCCAAAGCCAAAGGCACTAGTGTGTCTGATGATGATCTTCTAACCCGGAAGCTGGA 500
 Db 1199 TCCAAAGCCAAAGGCACTAGTGTGTCTGATGATGATCTTCTAACCCGGAAGCTGGA 1258
 QY 501 CAGTGCCCTGGAAGGCAATGCAAGCCCCCTCAAGGCGGAGTGAAGCAACCAACCTT 560
 Db 1259 CAGTGCCCTGGAAGGCAATGCAAGCCCCCTCAAGGCGGAGTGAAGCAACCAACCTT 1318
 QY 561 CCAAAGAGAGCAACAAGTACGGGCGCAGAGTACCTGAGCCTGAGGCCGAGAGCACT 620
 Db 1319 CCAAAGAGAGCAACAAGTACGGGCGCAGAGTACCTGAGCCTGAGGCCGAGAGCACT 1378
 QY 621 GGAAGTCCCAAGAACTACAGCTCCAGAGTCAAGCATGAAGGAGCAACCGTGAATGAG 680
 Db 1379 GGAAGTCCCAAGAACTACAGCTCCAGAGTCAAGCATGAAGGAGCAACCGTGAATGAG 1437
 QY 681 AAGAAGTGTCC 691
 Db 1438 AAGAAGTGTCC 1448

RESULT 12
 US-10-198-846-13529
 ; Sequence 13529, Application US/10198846
 ; Publication No. US2003009974A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lillie, James
 ; APPLICANT: Xu, Yongyao
 ; APPLICANT: Wang, Youzhen
 ; APPLICANT: Steinmann, Kathleen
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
 ; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ; FILE REFERENCE: MRI-049
 ; CURRENT APPLICATION NUMBER: US/10/198,846
 ; CURRENT FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: 60/306,220
 ; PRIOR FILING DATE: 2001-07-18
 ; NUMBER OF SEQ ID NOS: 14084
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13529
 ; LENGTH: 2635
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1520, 1599, 1663
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-198-846-13529

Query Match 34.6%; Score 505.8; DB 15; Length 2635;
 Best Local Similarity 86.9%; Pred. No. 8,8e-140;

Matches 592; Conservative 0; Mismatches 82; Indels 7; Gaps 3;
 QY 14 CTATTCCTACACCTCTCTCACTCAAGGACAGAGGTCTTGGGCCCAAGTCTGCTGATCAG 73
 Db 54 CTCTCCTCAACCTCTCTCACTCACTGACAGAGGTCTTGGGCCCAAGTCTGCTGATCAG 113
 QY 74 TCTGCTCCGCTCTGAGTCTCTCTGAGAGTCAATCAATCTCTGCACTGGAACACAG 133
 Db 114 CCAAGCTCAAGGCTCTGAGAGTCTCTCTGAGAGTCAATCAATCTCTGCACTGGAACACAG 173
 QY 134 AGTCAAGTGTGTGTATTAATCTATGCTCTCTGATCAACAGCAACCAAGCAAGCCCC 193
 Db 174 TCCAAAT---CGAAATATATATGTAAGTGTATCAAGAGTCTCCAGAAACGAGCCCC 230
 QY 194 AAATCATCATTTATGAGTCAAGTAATCGAGCCCTGAGGAGTTCTAATGCTTCTGAGC 253
 Db 231 AAATCTCTATCTATAGTAATATATGAGGCTCTGAGGAGTTCTGATCTGAGC 290
 QY 254 TCCAAAGTGTGCAACAGGCTCTCTGATCAATCTCTGAGGCTCTGAGGATGAGGCT 313
 Db 291 TCCAAAGTGTGCAACAGGCTCTCTGATCAATCTCTGAGGCTCTGAGGATGAGGCT 350
 QY 314 GATTATTAAGTGTGCTCAT---ATACAGAAATCTTCTCATGCTTCTGAACTGGAAC 370
 Db 351 GATTATTAAGTGTGCTCAT---ATACAGAAATCTTCTCATGCTTCTGAACTGGAAC 410
 QY 371 AAGTCAAGCTCTCTAGTCAAGCCCAAGGCAACCCCACTGATCTGTTCCGCGCTCC 430
 Db 411 AAGTCAAGCTCTCTAGTCAAGCCCAAGGCAACCCCACTGATCTGTTCCGCGCTCC 470
 QY 431 TCTGAGAGTCTCAAGCCCAAGGCAACCCCACTGATCTGTTCCGCGCTCC 490
 Db 471 TCTGAGAGTCTCAAGCCCAAGGCAACCCCACTGATCTGTTCCGCGCTCC 530
 QY 491 GGAAGTGTGCAAGTGTCTGAAAGGCAATGCAAGCCCCCTCAAGGCGGAGTGAAGACC 550
 Db 531 GGAAGTGTGCAAGTGTCTGAAAGGCAATGCAAGCCCCCTCAAGGCGGAGTGAAGACC 590
 QY 551 ACCAAACCTCTCAACAGAGCAACAAGTAAGCGGCGCAGAGCTACTGAGCTGAGC 610
 Db 591 ACCAAACCTCTCAACAGAGCAACAAGTAAGCGGCGCAGAGCTACTGAGCTGAGC 650
 QY 611 CCCGAGCAGTGAAGTCCCAAGAGCTACAGTCCAGAGTCAAGCATGAAGGAGAGACC 670
 Db 651 CCCGAGCAGTGAAGTCCCAAGAGCTACAGTCCAGAGTCAAGCATGAAGGAGAGACC 710
 QY 671 GTGATGAAGAGAGTGTCC 691
 Db 711 GTGATGAAGAGAGTGTCC 730

RESULT 13
 US-10-342-887-1913
 ; Sequence 1913, Application US/10342887
 ; Publication No. US20040058340A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dai, Hongyue
 ; APPLICANT: He, Yundong
 ; APPLICANT: Linsley, Peter S.
 ; APPLICANT: Mao, Mao
 ; APPLICANT: Roberts, Christopher J.
 ; APPLICANT: Van 't Veer, Laura Johanna
 ; APPLICANT: Van de Vijver, Marc J.
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; FILE REFERENCE: 9301-188-999
 ; CURRENT APPLICATION NUMBER: US/10/342,887
 ; CURRENT FILING DATE: 2003-01-15
 ; PRIOR APPLICATION NUMBER: 60/298,918
 ; PRIOR FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: 60/380,710
 ; PRIOR FILING DATE: 2002-05-14
 ; PRIOR APPLICATION NUMBER: 10/172,118


```
/ Publication No. US20020086303A1
/ GENERAL INFORMATION:
/ APPLICANT: Meagher, Madeleine Joy
/ APPLICANT: King, Gordon E.
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Secrist, Heather
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
/ FILE REFERENCE: 210121.532
/ CURRENT APPLICATION NUMBER: US/09/878,134
/ CURRENT FILING DATE: 2001-06-07
/ NUMBER OF SEQ ID NOS: 377
/ SOFTWARE: Corixa Invention Disclosure Database
/ SEQ ID NO 264
/ LENGTH: 605
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(605)
/ OTHER INFORMATION: n = A,T,C or G
US-09-878-134-264
```

```
Query Match 34.1%; Score 497.2; DB 13; Length 605;
Best Local Similarity 92.8%; Pred.No. 1.9e-137;
Matches 531; Conservative 0; Mismatches 40; Indels 1; Gaps 1;
```

```
QY 1 GGCCTGGAGCTGCTATTCTCTCACTCCCTCTCACTCAGGGGACAGGGGCTCGGGCCAGTC 60
DB 35 GGCCTGGAGCTGCTATTCTCTCACTCCCTCTCACTCAGGGGACAGGGGCTCGGGCCAGTC 94
QY 61 TGCCCTGACTAGTGCCTCCGCTCTGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 120
DB 95 TGCCCTGACTAGTGCCTCCGCTCTGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 154
QY 121 CACTGGAACCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 155 CACTGGAACCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 214
QY 181 AGGCAAGGCCCCCAACTCATCTATTATAGGTCAAGTATCGGCCCTCAGGGGTTCTTA 240
DB 215 AGGCAAGGCCCCCAACTCATCTATTATAGGTCAAGTATCGGCCCTCAGGGGTTCTTA 274
QY 241 TGGCTTCTTGGCTCCAGTCTGCGAACAAGGCTCTCCAGCATCTTGGGCTTCAGGC 300
DB 275 TGGCTTCTTGGCTCCAGTCTGCGAACAAGGCTCTCCAGCATCTTGGGCTTCAGGC 334
QY 301 TGAGGACGAGGCTGATTATTAAGTCTGCTCATATACAAAGTACTTCTCATGCTTCCG 360
DB 335 TGAGGACGAGGCTGATTATTAAGTCTGCTCATATACAAAGTACTTCTCATGCTTCCG 394
QY 361 AACTGGGACCAAGTCAAGCTCTAGGTCAAGGCCCAAGGCCCAAGGCCCAAGGCCCAAG 420
DB 395 CGAGAGGACCAAGTCAAGCTCTAGGTCAAGGCCCAAGGCCCAAGGCCCAAGGCCCAAG 454
QY 421 CCGGCTCTCTGAGGAGCTCCAGGCCAACAAGGCCAACAAGGCTCTGATCAAGTGA 480
DB 455 CCGGCTCTCTGAGGAGCTCCAGGCCAACAAGGCCAACAAGGCTCTGATCAAGTGA 514
QY 481 CTTCTACCGGAGAGCTGTGACAGTGGCTTGAAGGACAGATGGACAGCCCGTCAAGGCGG 540
DB 515 CTTCTACCGGAGAGCTGTGACAGTGGCTTGAAGGACAGATGGACAGCCCGTCAAGGCGG 573
QY 541 AGTGAAGACCAACCAACCTCTCAACAGAGCA 572
DB 574 AGTGAAGACCAACCAACCTCTCAACAGAGCA 605
```

Search completed: April 26, 2004, 22:54:12
Job time : 612.283 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 15:48:07 ; Search time 122.019 Seconds
(without alignments)
6640.180 Million cell updates/sec

Title: US-10-076-747-8

Perfect score: 1460

Sequence: 1 ggccttgagcctcgtcttc.....ggcagcagagccacgagct 1460

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 27747546 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/ina/5A COMB.seg:*
2: /cgn2_6/ptodata/2/ina/5B COMB.seg:*
3: /cgn2_6/ptodata/2/ina/6A COMB.seg:*
4: /cgn2_6/ptodata/2/ina/6B COMB.seg:*
5: /cgn2_6/ptodata/2/ina/PCUTUS COMB.seg:*
6: /cgn2_6/ptodata/2/ina/backfile1.seg:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	607	41.6	891	US-09-049-672A-23	Sequence 23, Appl
2	597.4	40.9	902	US-08-378-939-11	Sequence 11, Appl
3	498.4	34.1	935	US-09-049-672A-20	Sequence 20, Appl
4	480.6	32.9	543	US-09-702-705-970	Sequence 970, App
5	480.6	32.9	543	US-08-736-457-970	Sequence 970, App
6	480.6	32.9	543	US-09-614-124B-970	Sequence 970, App
7	480.6	32.9	543	US-09-671-325-970	Sequence 970, App
8	454.8	31.2	711	US-08-487-550-9	Sequence 9, Appl
9	454.8	31.2	711	US-09-526-098-9	Sequence 9, Appl
10	440.4	30.2	895	US-09-049-672A-25	Sequence 25, Appl
11	438.4	30.0	705	US-09-372-425A-3	Sequence 3, Appl
12	419.6	28.7	879	US-08-152-060-29	Sequence 29, Appl
13	418	28.6	885	US-09-152-060-47	Sequence 47, Appl
14	414	28.4	702	US-08-523-89A-5	Sequence 5, Appl
15	410.8	28.1	705	US-08-487-550-1	Sequence 1, Appl
16	410.8	28.1	705	US-09-526-098-1	Sequence 1, Appl
17	410.2	28.1	919	US-09-049-672A-24	Sequence 24, Appl
18	382.8	26.2	928	US-09-152-060-46	Sequence 46, Appl
19	379	26.0	771	US-08-991-789A-241	Sequence 241, App
20	379	26.0	771	US-09-062-451-241	Sequence 241, App
21	379	26.0	771	US-09-598-326-241	Sequence 241, App
22	379	26.0	771	US-09-289-198-241	Sequence 241, App
23	379	26.0	771	US-09-429-755-241	Sequence 241, App
24	337.6	23.1	491	US-09-702-705-833	Sequence 833, App
25	337.6	23.1	491	US-09-736-457-833	Sequence 833, App
26	337.6	23.1	491	US-09-614-124B-833	Sequence 833, App
27	337.6	23.1	491	US-09-671-325-833	Sequence 833, App

28	329.4	22.6	585	4	US-09-620-312D-551	Sequence 551, App
29	324.8	22.2	705	4	US-09-592-998C-7	Sequence 7, Appl
30	324.8	22.2	716	4	US-08-793-450-5	Sequence 5, Appl
31	323.4	22.2	608	4	US-09-702-705-908	Sequence 908, App
32	323.4	22.2	608	4	US-09-736-457-908	Sequence 908, App
33	323.4	22.2	608	4	US-09-614-124B-908	Sequence 908, App
34	323.4	22.2	608	4	US-09-671-325-908	Sequence 908, App
35	322.4	22.1	584	4	US-08-803-085-1	Sequence 1, Appl
36	322.4	22.1	584	4	US-09-404-879A-268	Sequence 268, App
37	322.4	22.1	584	4	US-09-338-933-268	Sequence 268, App
38	322.4	22.1	584	4	US-09-215-681-268	Sequence 268, App
39	322.4	22.1	584	4	US-09-215-681-268	Sequence 268, App
40	309	21.2	548	4	US-09-404-879A-267	Sequence 267, App
41	309	21.2	548	4	US-09-338-933-267	Sequence 267, App
42	309	21.2	548	4	US-09-215-681-267	Sequence 267, App
43	309	21.2	548	4	US-09-215-681-267	Sequence 267, App
44	301.4	20.6	735	4	US-10-039-785-63	Sequence 63, Appl
45	301.4	20.6	735	4	US-10-039-785-63	Sequence 64, Appl

ALIGNMENTS

RESULT 1
US-09-049-672A-23
Sequence 23, Application US/09049672A
Patent No. 6135941
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Tang, Y. Tom
APPLICANT: Yue, Henry
APPLICANT: Au-Yang, Janice
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: HUMAN IMMUNE SYSTEM ASSOCIATED PROTEINS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASTESQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/049,672A
FILING DATE: HERewith
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerrione, Michael C
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0497 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 891 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: THYMOTIO

CLONE: 2872705
US-09-049-672A-23

Query Match 41.6%; Score 607; DB 3; Length 891;
Best Local Similarity 93.3%; Pred. No. 6,8e-149;
Matches 645; Conservative 0; Mismatches 45; Indels 1; Gaps 1;

```

QY 1 GGCCTGGGCTGTGATTTCTTCACTCCTCCTCAGGAGCAAGGGCTCTGGGCCAATGTC 60
DB 36 GGCCTGGGCTGTGATTTCTTCACTCCTCCTCAGGAGCAAGGGCTCTGGGCCAATGTC 95
QY 61 TGCCTGACTCAGTCTGCTCCGTCTGGGCTCTCTGGAGAGTGCATCACCATTCTCTG 120
DB 96 TGCCTGACTCAGTCTGCTCCGTCTGGGCTCTCTGGAGAGTGCATCACCATTCTCTG 155
QY 121 CACTGGAACAGAGTCAAGTGGTGGTGAATTAATATGCTCTCTGGAGAGTGCATCACC 180
DB 156 CACTGGAACAGAGTCAAGTGGTGGTGAATTAATATGCTCTCTGGAGAGTGCATCACC 215
QY 181 AGGCAAGGCCCCCAACTCATTTATAGAGTCAAGTCAAGGCTCAGGGGTTTCTAA 240
DB 216 AGGCAAGGCCCCCAACTCATTTATAGAGTCAAGTCAAGGCTCAGGGGTTTCTAA 275
QY 241 TGCCTTCTTGGCTCCTCAAGTCTGGCAACAGGCTCCTGACATCTCTGGCTTCAAGC 300
DB 276 TCGGTTCTTGGCTCCTCAAGTCTGGCAACAGGCTCCTGACATCTCTGGCTTCAAGC 335
QY 301 TGAGACAGAGGCTGATTTATTAATGCTGCTCATATCAAGAGTCTTCTCATGCTTCCG 360
DB 336 TGAGACAGAGGCTGATTTATTAATGCTGCTCATATCAAGAGTCTTCTCATGCTTCCG 395
QY 361 AACTGGAGCAAGAGTCAAGTCTGCTGAGTCAAGGCAAGGCTCAGTCTGCTT 420
DB 396 CGAGAGGAGCAAGTCAAGTCTGCTGAGTCAAGGCAAGGCTCAGTCTGCTT 455
QY 421 CCGGCTCTCTGAGAGTCTCAAGGCAAGGCTCAGTCTGCTGAGTCAAGTCAAGTCA 480
DB 456 CCGGCTCTCTGAGAGTCTCAAGGCAAGGCTCAGTCTGCTGAGTCAAGTCAAGTCA 515
QY 481 CTTCTACCCGGAGGCTGAGAGTGGCTGAGAGGAGTGGAGGCTCCTGCAAGGCGG 540
DB 516 CTTCTACCCGGAGGCTGAGAGTGGCTGAGAGGAGTGGAGGCTCCTGCAAGGCGG 575
QY 541 AGTGAAGACCAACCAACCTTCCAAACAGAGCAACAAAGAGGCGGCAAGCTACT 600
DB 576 AGTGAAGACCAACCAACCTTCCAAACAGAGCAACAAAGAGGCGGCAAGCTACT 635
QY 601 GAGCTGAAGCGCCGAGAGTGAAGTCCACAGAGTCAAGTCAAGGCTCAAGCTGA 660
DB 636 GAGCTGAAGCGCCGAGAGTGAAGTCCACAGAGTCAAGTCAAGGCTCAAGCTGA 695
QY 661 AGGAGCAACCGTGAATGAAGACAGTGTCC 691
DB 696 AGGAGCAACCGTGAAGTGAAGACAGTGTCC 725

```

RESULT 2
US-08-378-939-11

Sequence 11, Application US/08378939

Patent No. 5876961

GENERAL INFORMATION:

APPLICANT: CROME, JAMES SCOTT

APPLICANT: LEWIS, ALAN PETER

TITLE OF INVENTION: PRODUCTION OF ANTIBODIES

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ

STREET: 555 THIRTEENTH ST. N.W.

CITY: WASHINGTON

STATE: D. C.

COUNTRY: U.S.

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/378,939
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/952640
FILING DATE: 01-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: ERNST, BARBARA G
REGISTRATION NUMBER: 30,377
REFERENCE/DOCKET NUMBER: 1808-118
TELEPHONE: (202) 783-6040
TELEFAX: (202) 783-6031
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 902 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 32..739
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 89..739
FEATURE:
NAME/KEY: 91g_peptide
LOCATION: 32..86
US-08-378-939-11

Query Match 40.9%; Score 597.4; DB 2; Length 902;
Best Local Similarity 92.5%; Pred. No. 2.2e-146;
Matches 639; Conservative 0; Mismatches 51; Indels 1; Gaps 1;

```

QY 1 GGCCTGGGCTGTGATTTCTTCACTCCTCCTCAGGAGCAAGGGCTCTGGGCCAATGTC 60
DB 34 GGCCTGGGCTGTGATTTCTTCACTCCTCCTCAGGAGCAAGGGCTCTGGGCCAATGTC 93
QY 61 TGCCTGACTCAGTCTGCTCCGTCTGGGCTCTCTGGAGAGTGCATCACCATTCTCTG 120
DB 94 TGCCTGACTCAGTCTGCTCCGTCTGGGCTCTCTGGAGAGTGCATCACCATTCTCTG 153
QY 121 CACTGGAACAGAGTCAAGTGGTGGTGAATTAATATGCTCTCTGGAGAGTGCATCACC 180
DB 154 CACTGGAACAGAGTCAAGTGGTGGTGAATTAATATGCTCTCTGGAGAGTGCATCACC 213
QY 181 AGGCAAGGCCCCCAACTCATTTATAGAGTCAAGTCAAGGCTCAGGGGTTTCTAA 240
DB 214 AGGCAAGGCCCCCAACTCATTTATAGAGTCAAGTCAAGGCTCAGGGGTTTCTAA 273
QY 241 TGCCTTCTTGGCTCCTCAAGTCTGGCAACAGGCTCCTGACATCTCTGGGCTTCAAGC 300
DB 274 TGCCTTCTTGGCTCCTCAAGTCTGGCAACAGGCTCCTGACATCTCTGGGCTTCAAGC 333
QY 301 TGAGACAGAGGCTGATTTATTAATGCTGCTCATATCAAGAGTCAAGTCAAGTCTT 360
DB 334 TGAGACAGAGGCTGATTTATTAATGCTGCTCATATCAAGAGTCAAGTCAAGTCTT 393
QY 361 AACTGGAGCAAGAGTCAAGTCTGAGTCAAGGCAAGGCTCAGGCTCAAGTCTGCT 420
DB 394 CGAGAGGAGCAAGAGTCAAGTCTGAGTCAAGGCAAGGCTCAGGCTCAAGTCTGCT 453
QY 421 CCGGCTCTCTTGAAGAGTCTCAAGCAACAGGCTCAAGTCTGAGTCTGATCAAGTGA 480
DB 454 CCGGCTCTCTTGAAGAGTCTCAAGCAACAGGCTCAAGTCTGAGTCTGATCAAGTGA 513

```

QY 481 CTTCTACCGGAGCTGTGACAGTGGCTTGGAGGAGATGGCGCCCGTCAAGCGGG 540
Db 514 CTTCTACCGGAGCGGTGTGACAGTGGCTTGGAGGAGATGGCGCCCGTCAAGCGGG 573
QY 541 AGTGGAGACCAACCAACCCCTCCCAACAGAGCAACCAAGTAGCGGGCCAGCACTACT 600
Db 574 AGTGGAGACCAACCAACCCCTCCCAACAGAGCAACCAAGTAGCGGGCCAGCACTACT 633
QY 601 GAGCTGACCGCCGAGCAGTGAAGTCCACAGAGCTACAGCTGCGCAGTCAGCATGA 660
Db 634 GAGCTGACCGCCGAGCAGTGAAGTCCACAGAGCTACAGCTGCGCAGTCAGCATGA 693
QY 661 AGGAGACCGCTGATGAAGAACAAGTGCTCC 691
Db 694 AGGAGACCGCTGGA-GAAGACAGTGCCCC 723

RESULT 3

US-09-049-672A-20
Sequence 20, Application US/09049672A
Patent No. 6135941
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Tang, Y. Tom
APPLICANT: Yue, Henry
APPLICANT: Au-Young, Janice
APPLICANT: Corley, Neil C.
APPLICANT: Guebler, Karl J.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: HUMAN IMMUNE SYSTEM ASSOCIATED PROTEINS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/049,672A
FILING DATE: HEREWITH
CLASSIFICATION: 536
PRIOR APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Carrone, Michael C
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0497 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 935 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: ADREUT05
CLONE: 2492122
US-09-049-672A-20

Query Match 34.1%; Score 498.4; DB 3; Length 935;
Best Local Similarity 81.6%; Pred. No. 1,4e-120;

Matches 627; Conservative 0; Mismatches 131; Indels 10; Gaps 4;
QY 1 GAGCTGGGCTCTGCTATTTCCTACCCCTGCTACTACAGGGCAAGAGTCTGGGCGCAGTC 60
Db 76 GAGCTGGTCT 135
QY 61 TGCCCTGACTAGTCTGCT 120
Db 136 TGCTGTGACGACGCGCCCTCAGTGTCTGGGGCCCAAGGAGAGGGGTCAACATCTCTG 195
QY 121 CACTGGAAACGACAGTACAGTGTGTATTAATATGTCTCTCTGTATACCAAGCAACC 180
Db 196 CACTGGAGACGCTCCCAACATGGGGCAAGTTAATGATACCTGTATCAAGCAAGCTTCC 255
QY 181 AGGCAAGCCCGCCAACTCATTTATGAAGTCAATGAGGCTCAGGGTTCTTAA 240
Db 256 AGGAACAGCCCGCCAACTCATTTATGAAGTCAATGAGGCTCAGGGTTCTTAA 315
QY 241 TGCT 300
Db 316 CCGATTCTCTGCTCCAAAGTCTGCACTCAGCTCTCTCTGCGCATCACTGGGCTCCAGGC 375
QY 301 TGAGGACGAGGCTGATTATTACTGCTGCTCATATACAAAG--TACTTCTCATGCTT 357
Db 376 TGAGGATGAGGCTGATTATTACTGCTGCTCATATACAGGAGCTGAGTGTGATT 435
QY 358 CGGAAGTGGACCAAGGTCACCGTCTTAAAGTCAAGCCCAAGCCCACTGCTCACTCT 417
Db 436 CGGCGAGAGGACCAAGGTCACCGTCTTAAAGTCAAGCCCAAGCCCACTGCTGCTCACTCT 495
QY 418 GTTCCCGCCCT 477
Db 496 GTTCCCGCCCT 555
QY 478 TGACTTCAACCGGAGAGCTGTGAACAGTGGCTGGAAGGAGATGGCAAGCCCGTCAAGGC 537
Db 556 TGACTTCAACCGGAGAGCTGTGAACAGTGGCTGGAAGGAGATGGCAAGCCCGTCAAGGC 615
QY 538 GGGAGTGAAGACCAACCACTTCCAAACAGAGCAACCAAGTACGGGCGAGCACTA 597
Db 616 GGGAGTGAAGACCAACCACTTCCAAACAGAGCAACCAAGTACGGGCGAGCACTA 675
QY 598 CCGAGGCTGACCGCGAGAGTGAAGTCCACAGAGTACAGCTGCGCAGGTCAAGCA 657
Db 676 CCGAGGCTGACCGCGAGAGTGAAGTCCACAGAGTACAGCTGCGCAGGTCAAGCA 735
QY 658 TGAAGGAGACCGGTGATGAAGACAGTGTCTCCCTACAGAAATGTTCTATAGTTCTCT 717
Db 736 TGAAGGAGACCGGTGATGAAGACAGTGTCTCCCTACAGAAATGTTCTATAGTTCTCT 789
QY 718 AACAGCTCAAGCCCGCACTACCAATGGCGAGACTAGCAGCTGCAAGG 765
Db 790 -AACAGCTCAAGCCCGCACTACCAAGTGAAGCTGAGATGCCAGG 836

RESULT 4

US-09-702-705-970/c
Sequence 970, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaltanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705

;; CURRENT FILING DATE: 2000-10-30
;; NUMBER OF SEQ ID NOS: 1833
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 970
;; LENGTH: 543
;; TYPE: DNA
;; ORGANISM: Homo sapien
US-09-702-705-970

Query Match 32.9%; Score 480.6; DB 4; Length 543;
Best Local Similarity 92.8%; Pred. No. 5,3e-116;
Matches 504; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 100 ACAGTCGATCACCATTCTCTGCACTGGAACGACGAGTCAGTGGTGTATTAATATGT 159
DB 543 ACAGTCGATCACCATTCTCTGCACTGGAACGACGAGTCAGTGGTGTATTAATATGT 484
QY 160 CTCCTGGTACCAACAGACACCCAGGAAAGCCCCCAATCATCTATTATGAGGTCACTAA 219
DB 483 CTCCTGGTACCAACAGACACCCAGGAAAGCCCCCAATCATCTATTATGATGTCACTGA 424
QY 220 TCGGACCTCAGGGGTTTCTTAATCGCTCTCTGCTCCAGTCTGCAACACGAGCTCCCT 279
DB 423 TCGGACCTCAGGGGTTTCTTAATCGCTCTCTGCTCCAGTCTGCAACACGAGCTCCCT 364
QY 280 GACCATCTCTGGGCTCCAGGCTGAGAGACGAGGCTGATTATTACTGCTGCTCATATACAG 339
DB 363 GACCATCTCTGGGCTCCAGGCTGAGAGACGAGGCTGATTATTACTGCACTCATATACAA 304
QY 340 AAGTACTTCTCATGTCTTCCGGAACGAGGTCACGCTCCTAGTCAAGCCCAAGGC 399
DB 303 CAGTACGACTTCTGGTGTTCGGAGGAGGACCAAGGCTGACCTCTAGGTCAAGCCCAAGGC 244
QY 400 CAACCCCACTGCTCACTGTTCCTCCGCTCTCTGAGAGGCTCCAGACCAACAGGCTCAC 459
DB 243 TCGCCCTCTGCTCACTGTTCCTCCGCTCTCTGAGAGGCTCCAGACCAACAGGCTCAC 184
QY 460 ACTAGTGTCTGATCAGTGAATTTACCCGAGAGCTGTGACAGTGGCTTGAAGGCGAGA 519
DB 183 ACTAGTGTCTGATCAGTGAATTTACCCGAGAGCTGTGACAGTGGCTTGAAGGCGAGA 124
QY 520 TGGCAGCCCCCGTCAAGGGGGAGTGAAGACCAACCCCTCCAAACAGAGCAACAA 579
DB 123 TAGCAGCCCCCGTCAAGGGGGAGTGAAGACCAACCCCTCCAAACAGAGCAACAA 64
QY 580 GTACCGGGCCAGACAGTACTGAGCTTGAGCGCCGAGCAGTGAAGTCCACAGAACTA 639
DB 63 GTACCGGGCCAGACAGTACTGAGCTTGAGCGCCGAGCAGTGAAGTCCACAGAACTA 4
QY 640 CAG 642
DB 3 CAG 1

RESULT 5

US-09-736-457-970/c
; Sequence 970, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Veddyck, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fat, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478CIS
; CURRENT APPLICATION NUMBER: US/09/736,457

;; CURRENT FILING DATE: 2000-12-13
;; NUMBER OF SEQ ID NOS: 1864
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 970
;; LENGTH: 543
;; TYPE: DNA
;; ORGANISM: Homo sapien
US-09-736-457-970

Query Match 32.9%; Score 480.6; DB 4; Length 543;
Best Local Similarity 92.8%; Pred. No. 5,3e-116;
Matches 504; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 100 ACAGTCGATCACCATTCTCTGCACTGGAACGACGAGTCAGTGGTGTATTAATATGT 159
DB 543 ACAGTCGATCACCATTCTCTGCACTGGAACGACGAGTCAGTGGTGTATTAATATGT 484
QY 160 CTCCTGGTACCAACAGACACCCAGGAAAGCCCCCAATCATCTATTATGAGGTCACTAA 219
DB 483 CTCCTGGTACCAACAGACACCCAGGAAAGCCCCCAATCATCTATTATGATGTCACTGA 424
QY 220 TCGGACCTCAGGGGTTTCTTAATCGCTCTCTGCTCCAGTCTGCAACACGAGCTCCCT 279
DB 423 TCGGACCTCAGGGGTTTCTTAATCGCTCTCTGCTCCAGTCTGCAACACGAGCTCCCT 364
QY 280 GACCATCTCTGGGCTCCAGGCTGAGAGACGAGGCTGATTATTACTGCTGCTCATATACAG 339
DB 363 GACCATCTCTGGGCTCCAGGCTGAGAGACGAGGCTGATTATTACTGCACTCATATACAA 304
QY 340 AAGTACTTCTCATGTCTTCCGGAACGAGGTCACGCTCCTAGTCAAGCCCAAGGC 399
DB 303 CAGTACGACTTCTGGTGTTCGGAGGAGGACCAAGGCTGACCTCTAGGTCAAGCCCAAGGC 244
QY 400 CAACCCCACTGCTCACTGTTCCTCCGCTCTCTGAGAGGCTCCAGACCAACAGGCTCAC 459
DB 243 TCGCCCTCTGCTCACTGTTCCTCCGCTCTCTGAGAGGCTCCAGACCAACAGGCTCAC 184
QY 460 ACTAGTGTCTGATCAGTGAATTTACCCGAGAGCTGTGACAGTGGCTTGAAGGCGAGA 519
DB 183 ACTAGTGTCTGATCAGTGAATTTACCCGAGAGCTGTGACAGTGGCTTGAAGGCGAGA 124
QY 520 TGGCAGCCCCCGTCAAGGGGGAGTGAAGACCAACCCCTCCAAACAGAGCAACAA 579
DB 123 TAGCAGCCCCCGTCAAGGGGGAGTGAAGACCAACCCCTCCAAACAGAGCAACAA 64
QY 580 GTACCGGGCCAGACAGTACTGAGCTTGAGCGCCGAGCAGTGAAGTCCACAGAACTA 639
DB 63 GTACCGGGCCAGACAGTACTGAGCTTGAGCGCCGAGCAGTGAAGTCCACAGAACTA 4
QY 640 CAG 642
DB 3 CAG 1

RESULT 6

US-09-614-124B-970/c
; Sequence 970, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Veddyck, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fat, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668

SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 970
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-970

Query Match 32.9%; Score 480.6; DB 4; Length 543;
Best Local Similarity 92.8%; Pred. No. 5.3e-116;
Matches 504; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```
QY 100 ACAGTCATGACCATCTCTGCACTGAGACAGAGTCACTGGTGGTTATATCATATGT 159
DB 543 ACAGTCATGACCATCTCTGCACTGAGACAGAGTCACTGGTGGTTATATCATATGT 484
QY 160 CTCCTGTACCAACAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 219
DB 483 CTCCTGTACCAACAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 424
QY 220 TCGGCCCCAGAGGGTTTCTAATCGCTTCTGCTCTCAAGTCTGGCAACAGGCTCCCT 279
DB 423 TCGGCCCCAGAGGGTTTCTAATCGCTTCTGCTCTCAAGTCTGGCAACAGGCTCCCT 364
QY 280 GACCATCTGAGGCTCCAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 339
DB 363 GACCATCTGAGGCTCCAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 304
QY 340 AAGTACTTCTCATGCTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 399
DB 303 AAGTACTTCTCATGCTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 244
QY 400 CAACCCCACTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459
DB 243 TCGCCCCAGAGGGTTTCTAATCGCTTCTGCTCTCAAGTCTGGCAACAGGCTCCCT 184
QY 460 ACTAGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 519
DB 183 ACTAGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 124
QY 520 TGGCAGCCCCGTCAAGCGGGAGTGGAGACACCAACCTCCAAACAGAGCAACA 579
DB 123 TGGCAGCCCCGTCAAGCGGGAGTGGAGACACCAACCTCCAAACAGAGCAACA 64
QY 580 GTACGGCGGAGAGAGTCACTGAGCTGAGCGCCGAGAGAGAGAGAGAGAGAG 639
DB 63 GTACGGCGGAGAGAGTCACTGAGCTGAGCGCTGAGCGAGTGGAGAGTCCCAAAAGCTA 4
QY 640 CAG 642
DB 3 CAG 1
```

RESULT 7

US-09-671-325-970/c
; Sequence 970, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongrong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Pan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 970
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-970

Query Match 32.9%; Score 480.6; DB 4; Length 543;
Best Local Similarity 92.8%; Pred. No. 5.3e-116;
Matches 504; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```
QY 100 ACAGTCATGACCATCTCTGCACTGAGACAGAGTCACTGGTGGTTATATCATATGT 159
DB 543 ACAGTCATGACCATCTCTGCACTGAGACAGAGTCACTGGTGGTTATATCATATGT 484
QY 160 CTCCTGTACCAACAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 219
DB 483 CTCCTGTACCAACAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 424
QY 220 TCGGCCCCAGAGGGTTTCTAATCGCTTCTGCTCTCAAGTCTGGCAACAGGCTCCCT 279
DB 423 TCGGCCCCAGAGGGTTTCTAATCGCTTCTGCTCTCAAGTCTGGCAACAGGCTCCCT 364
QY 280 GACCATCTGAGGCTCCAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 339
DB 363 GACCATCTGAGGCTCCAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 304
QY 340 AAGTACTTCTCATGCTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 399
DB 303 AAGTACTTCTCATGCTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 244
QY 400 CAACCCCACTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459
DB 243 TCGCCCCAGAGGGTTTCTAATCGCTTCTGCTCTCAAGTCTGGCAACAGGCTCCCT 184
QY 460 ACTAGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 519
DB 183 ACTAGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 124
QY 520 TGGCAGCCCCGTCAAGCGGGAGTGGAGACACCAACCTCCAAACAGAGCAACA 579
DB 123 TGGCAGCCCCGTCAAGCGGGAGTGGAGACACCAACCTCCAAACAGAGCAACA 64
QY 580 GTACGGCGGAGAGAGTCACTGAGCTGAGCGCCGAGAGAGAGAGAGAGAGAG 639
DB 63 GTACGGCGGAGAGAGTCACTGAGCTGAGCGCTGAGCGAGTGGAGAGTCCCAAAAGCTA 4
QY 640 CAG 642
DB 3 CAG 1
```

RESULT 8

US-08-487-550-9
; Sequence 9, Application US/08487550
; Patent No. 6113898
; GENERAL INFORMATION:
; APPLICANT: Anderson, Darrell R.
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES SPECIFIC
; TITLE OF INVENTION: TO HUMAN B7.1 AND/OR B7.2 PRIMATED FORMS THEREOF,
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS
; TITLE OF INVENTION: IMMUNOSUPPRESSANTS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,550
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 012712-131
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: CDS
LOCATION: 1..711
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..711
US-08-487-550-9

Query Match 31.2%; Score 454.8; DB 3; Length 711;

Best Local Similarity 81.4%; Pred. No. 3.1e-109; Matches 565; Conservative 0; Mismatches 122; Indels 7; Gaps 3;

1 GGCCTGGGCTCTGCTATTCCTCACTCAGGAGGAGGCTCGGAGCCAGTC 60
6 GGTCCCGCTCAGCTCTGGGGCTCTGCTGCTCCAGGTGACAGATGAGATC 65
61 TGCCTGACTGACTGCTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
66 TGTCTGACACAGCCCGCTCAGTGTCTGAGGAGGAGGAGGAGGAGGAGGAG 125
121 CACTGAGACGAGGAGTCAAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCT 180
126 CACTGAGACGAGTCAAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCTGCT 182
181 AGGCAAGAGCCCGCTCAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCTGCT 240
183 AGGCAAGAGCCCGCTCAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCTGCT 242
241 TGCCTGACTGACTGCTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
243 CCGATTCCTGCTGCTCAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCTGCT 302
301 TGAAGACGAGGAGTCAAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCTGCT 357
303 TGAAGACGAGGAGTCAAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCTGCT 362
358 CGGAACTGAGGAGTCAAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCTGCT 417
363 CGGAGAGGAGTCAAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCTGCTGCT 422
418 GTTCCCGCTCAGCTCTGAGGAGTCAAGTGTGAGTATTAATGCTCTGCTGCTGCT 477
423 GTTCCCGCTCAGCTCTGAGGAGTCAAGTGTGAGTATTAATGCTCTGCTGCTGCT 482
478 TGAAGTCAAGGAGTCAAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCTGCT 537
483 TGAAGTCAAGGAGTCAAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCTGCT 542
538 GAGAGTCAAGGAGTCAAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCTGCT 597
543 GAGAGTCAAGGAGTCAAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCTGCT 602
598 CTTGAGTCAAGGAGTCAAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCTGCT 657

DB 603 CTTGAGTCAAGGAGTCAAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCTGCT 662
CY 658 TGAAGAGAGGAGTCAAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCTGCT 691
DB 663 TGAAGAGAGGAGTCAAGTGTGAGTATTAATGCTCTGCTGCTGCTGCTGCTGCT 695

RESULT 9
US-09-526-098-9
Sequence 9, Application US/09526098

Patent No. 6492134

GENERAL INFORMATION:

APPLICANT: Anderson, Darrell R.

TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC

TITLE OF INVENTION: TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF,

TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS

STREET: 699 Prince Street

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22314

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/526,098

FILING DATE:

CLASSIFICATION:

Prior Application Number: 09/383,916

FILING DATE:

APPLICATION NUMBER: US 08/487,550

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Teskin, Robin L.

REGISTRATION NUMBER: 35,030

REFERENCE/DOCKET NUMBER: 012712-131

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-836-6620

TELEFAX: 703-836-2021

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 711 base pairs

TYPE: nucleic acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: CDS

LOCATION: 1..711

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 1..711

US-09-526-098-9

Query Match 31.2%; Score 454.8; DB 4; Length 711;

Best Local Similarity 81.4%; Pred. No. 3.1e-109; Matches 565; Conservative 0; Mismatches 122; Indels 7; Gaps 3;

1 GGCCTGGGCTCTGCTATTCCTCACTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
6 GGTCCCGCTCAGCTCTGGGGCTCTGCTGCTCCAGGTGACAGATGAGATC 65
61 TGCCTGACTGACTGCTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
66 TGTCTGACACAGCCCGCTCAGTGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 125

QY 121 CACTGGAAACGACGATCAGCTTGGTGTATTAATATGTCTCTGTTACCAACGACCC 180
 Db 126 CACTGGAGGACCTCCCAACATGTAGGTTAT--GATCTACATTGGTACAGACGCTCCC 182
 QY 181 AGGCAAAAGCCCCCAATCATATTATAGAGTCAGTAAATCGGCCCCCAGGGTTCTAA 240
 Db 183 AGGAAGGCCCCCAACCTCATCTATGACATTAACAGAGAACCTCAAGAAATTTCTGA 242
 QY 241 TCGCTTCTGTGGCTTCAAGTCTGGCAACAGGCTCTCCCTGACCATCTCTGGCTCCAGG 300
 Db 243 CGGATTTCTGTGGCTTCAAGTCTGGCAACAGGCTCTCCCTGACCATCTCTGGCTCCAGG 302
 QY 301 TGAGGACGAGGTGATTTATTAATGCTGTCTCATATACAAAG--TACTTCTCATGTCTT 357
 Db 303 TGAGGATAGGCTGATTTATTAATGCTGTCTCATATGACAGAGCTGATGTCTCATGAT 362
 QY 358 CGGAACGTGGACCAAGGTACCGTCTTCAAGTCAAGCCCAAGGCCAACCCCACTGTACTCT 417
 Db 363 CGGAGAGGAGGACCGGCTGACCGTCTTCAAGTCAAGCCCAAGGCCAACCCCACTGTACTCT 422
 QY 418 GTTCCCGGCTCTCTGAGAGCTCCCAAGCCCAAGGCCCACTAGTGTCTGATCAG 477
 Db 423 GTTCCCGGCTCTCTGAGAGCTCCCAAGCCCAAGGCCCACTAGTGTCTGATCAG 482
 QY 478 TGACTTCTAACCCGAGAGCTGTGACAGTGGCTTGAAGGAGATGGACGCCCTCAAGGC 537
 Db 483 TGACTTCTAACCCGAGAGCTGTGACAGTGGCTTGAAGGAGATGGACGCCCTCAAGGC 542
 QY 538 GGGAGTGGAGACCAACCAACCTCTCCAAACAGAGCAACAAGTACGGGGCCAGACCTA 597
 Db 543 GGGAGTGGAGACCAACCAACCTCTCCAAACAGAGCAACAAGTACGGGGCCAGACCTA 602
 QY 598 CCTGAGCTTGAAGCCCGGAGAGTGGAGTCCCAAGAAAGTACAGCTGCGAGGTCAAGCA 657
 Db 603 CCTGAGCTTGAAGCCCGGAGAGTGGAGTCCCAAGAAAGTACAGCTGCGAGGTCAAGCA 662
 QY 658 TGAAGGAGACACCGTGTGATGAAGACAGTGTGCC 691
 Db 663 TGAAGGAGACACCGTGTGATGAAGACAGTGTGCC 695

RESULT 10
 US-09-049-672A-25
 ; Sequence 25, Application US/09049672A
 ; Patent No. 6135941
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Lal, Preeti
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Yue, Henry
 ; APPLICANT: Au-Young, Janice
 ; APPLICANT: Corley, Neil C.
 ; APPLICANT: Guegler, Karl J.
 ; APPLICANT: Baughn, Mariah R.
 ; TITLE OF INVENTION: HUMAN IMMUNE SYSTEM ASSOCIATED PROTEINS
 ; NUMBER OF SEQUENCES: 28
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/049,672A
 ; FILING DATE: HEREWITH
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:

APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Cerrione, Michael C
 REGISTRATION NUMBER: 39,132
 REFERENCE/DOCKET NUMBER: PF-0497 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-855-0555
 TELEFAX: 650-845-4166
 TELEX:
 INFORMATION FOR SEQ ID NO: 25:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 895 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: LUNGUT13
 CLONE: 3116314
 US-09-049-672A-25
 Query Match 30.2%; Score 440.4; DB 3; Length 895;
 Best Local Similarity 80.1%; Pred. No. 1.9e-105;
 Matches 556; Conservative 0; Mismatches 131; Indels 7; Gaps 3;
 QY 1 GGGCTGGGCTGTGCTATTCTTCAACCTCTCACTCAGGAGCAGAGGCTCTGGGCCCCAGTC 60
 Db 39 GGGCTGGGCTGTGCTATTCTTCAACCTCTCACTCAGGAGCAGAGGCTCTGGGCCCCAGTC 98
 QY 61 TGGCTGACTCAGTGTGCTCTGCTGTCTGAGTCTGAGTCACTGATGATCACTCTCTG 120
 Db 99 TGGCTGACTCAGTGTGCTCTGCTGTCTGAGTCTGAGTCACTGATGATCACTCTCTG 158
 QY 121 CACTGGAACGACAGTCAAGTGTGATGATTAATGATGCTCTGATGATCAACAGACACC 180
 Db 159 TTTTGAACCACTCTCAACATGCAAG--TAATCTGTGATGATGATGATGATGATGATGAT 215
 QY 181 AGGCAAAAGCCCCCAACCTCATATTATGAGTCAAGTATGGCCCTCAAGGATTTCTAA 240
 Db 216 AGGAGGAGGCCCCCAACCTCTCATATGCTATGATGATGATGATGATGATGATGATGAT 275
 QY 241 TGGCTTCTGTGCTCTCAAGTCTGAGCAACAGGCTCTGAGCACTCTGAGGCTCCAGGC 300
 Db 276 CGGATTTCTGTGCTCTCAAGTCTGAGCAACAGGCTCTGAGCACTCTGAGGCTCCAGGC 335
 QY 301 TGAGGACGAGGCTGATTAATTAATGCTGTCTCAT--ATACAAAGTACTTCTCATGTCTT 357
 Db 336 CGAGGATGAAGTGAATTAATTAATGCTGTCTCAT--ATACAAAGTACTTCTCATGTCTT 395
 QY 358 CGGAACGTGGAGACCAAGTCAAGTCTCTGAGTCAAGCCCAAGGCCCACTGTCACTCT 417
 Db 396 CGGAGGAGGAGCAAGTCAAGTCTCTGAGTCAAGCCCAAGGCCCACTGTCTCACTCT 455
 QY 418 GTTCCCGGCTCTCTGAGAGCTCCAAAGCCAAAGGCCCACTAGTGTCTGATCAG 477
 Db 456 GTTCCCGGCTCTCTGAGAGCTCCAAAGCCCAAGGCCCACTAGTGTCTGATCAG 515
 QY 478 TGACTTTAACCAGGAGCTGTGACAGTGGCTTGAAGGAGTGAAGGAGGCTCCGTCAAGG 537
 Db 516 TGACTTTAACCAGGAGCTGTGACAGTGGCTTGAAGGAGTGAAGGAGGCTCCGTCAAGG 575
 QY 538 GGGAGTGGAGACCAACCAACCTCTCCAAACAGAGCAACAAGTACGGGGCCAGAGCTA 597
 Db 576 GGGAGTGGAGACCAACCAACCTCTCCAAACAGAGCAACAAGTACGGGGCCAGAGCTA 635
 QY 598 CCGAGCTTGAAGCCCGAGCAAGTGAAGTCCCAAGAAAGTCAAGTGTGAGTCAAGCA 657
 Db 636 CCGAGCTTGAAGCCCGAGCAAGTGAAGTCCCAAGAAAGTCAAGTGTGAGTCAAGCA 695
 QY 658 TGAAGGAGACACCGTGTGATGAAGACAGTGTGCC 691
 Db 696 TGAAGGAGACACCGTGTGATGAAGACAGTGTGCC 728

RESULT 11

US-09-372-425A-3
 / Sequence 3, Application US/09372425A
 / Patent No. 6475749
 / GENERAL INFORMATION:
 / APPLICANT: Sherie L. Morrison
 / APPLICANT: Ramon Montano
 / TITLE OF INVENTION: Improved Rh Antibody
 / NUMBER OF SEQUENCES: 11
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
 / STREET: 2029 Century Park East, Suite 3800
 / CITY: Los Angeles
 / STATE: CA
 / COUNTRY: USA
 / ZIP: 90067
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy Disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: Windows 98
 / SOFTWARE: MS Word
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/09/372,425A
 / FILING DATE: August 11, 1999
 / CLASSIFICATION: 435
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER:
 / FILING DATE:
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Oldenkamp, David J.
 / REGISTRATION NUMBER: 29,421
 / REFERENCE/DOCKET NUMBER: 510015-223
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (310) 768-5100
 / TELEFAX: (310) 768-5100
 / INFORMATION FOR SEQ ID NO: 3:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 705 nucleotides
 / TYPE: nucleotide
 / STRANDEDNESS: single
 / TOPOLOGY: linear
 / MOLECULE TYPE: Light chain - DNA
 / US-09-372-425A-3

Query Match 30.0%; Score 438.4; DB 4; Length 705;
 Best Local Similarity 82.6%; Pred. No. 5.8e-105;
 Matches 565; Conservative 0; Mismatches 106; Indels 13; Gaps 5;

14 CTATTCCTCAACCTCTCTCACTCAGGGACAGGGTCTTGAGCCCACTTGGCCCTGACTCAG 73
 13 CTACTCTCAACCTCTCTCACTCAGGGTCTTGAGCCCACTTGGCCCTGACTCAG 72
 74 TCGGCTCGGTGTCTGAGTCTCTGAGAGTGCATCACCATCTCTGCACTGGAACAGC 133
 73 CCGCCCTCACTGTCTGAGTCTCTGAGAGTGCATCACCATCTCTGCACTGGAACAGC 132
 134 AGTCAGTGTGTGTATTAATATGTCTCTGTGTACCAAGACCCGGAAGCCCC 193
 133 TCCAAAT--TGGGAATATTAATATGTATCTGTGTACCAAGATTTCCGGAACACCC 189
 194 AAACATCATATTATAGGTAGTATGAGGCTTCTAATGAGTCTCTCTGAG 253
 190 AAATCTCTATTATATACATATATGAGGCTTCTAATGAGTCTCTCTGAG 249
 254 TCCAACTGGAACAACGAGCTCTCTGAGGCTTCTGAGGCTTCTGAGGCTTCTGAG 313
 250 TCCAACTGGAACAACGAGCTCTCTGAGGCTTCTGAGGCTTCTGAGGCTTCTGAG 309
 314 GATTATTAATGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 370
 310 GATTATTAATGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 369

QY 371 AAGTCAACGCTCTAGGTGAGCCCAAGGCCAACCCCACTGTCACTGTTC---CGGCC 427
 DB 370 GAATGACGCTCTAGGTGAGCCCAAGGCCAACCCCACTGTCTGTTC---CGGCC 429
 QY 428 TCTCTGAGAGGCTTCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 487
 DB 430 TCTCTGAGAGGCTTCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 489
 QY 488 CCGGAGCTGTGACAGTGTCTGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 547
 DB 490 CCGGAGCTGTGACAGTGTCTGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 549
 QY 548 ACCAACCAACCTTCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 607
 DB 550 ACCAACCAACCTTCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 606
 QY 608 ACCGCGGAGGAGTGAAGTCTTCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 667
 DB 607 ACCGCGGAGGAGTGAAGTCTTCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 666
 QY 668 ACCGTGATGAAGACAGTGTCTCC 691
 DB 667 ACCGTGATGAAGACAGTGTCTCC 689

RESULT 12

US-09-152-060-29
 / Sequence 29, Application US/09152060
 / Patent No. 6448230
 / GENERAL INFORMATION:
 / APPLICANT: Rosen et al.
 / TITLE OF INVENTION: 28 Human Secreted Proteins
 / FILE REFERENCE: P2003PL US
 / CURRENT APPLICATION NUMBER: US/09/152,060
 / FILING DATE: 1998-09-11
 / EARLIER APPLICATION NUMBER: PCT/US98/04858
 / EARLIER FILING DATE: 1998-03-12
 / EARLIER APPLICATION NUMBER: 60/040,762
 / EARLIER FILING DATE: 1997-03-14
 / EARLIER APPLICATION NUMBER: 60/040,710
 / EARLIER FILING DATE: 1997-03-14
 / EARLIER APPLICATION NUMBER: 60/050,934
 / EARLIER FILING DATE: 1997-05-30
 / EARLIER APPLICATION NUMBER: 60/048,100
 / EARLIER FILING DATE: 1997-05-30
 / EARLIER APPLICATION NUMBER: 60/048,357
 / EARLIER FILING DATE: 1997-05-30
 / EARLIER APPLICATION NUMBER: 60/048,189
 / EARLIER FILING DATE: 1997-05-30
 / EARLIER APPLICATION NUMBER: 60/057,765
 / EARLIER FILING DATE: 1997-09-05
 / EARLIER APPLICATION NUMBER: 60/048,970
 / EARLIER FILING DATE: 1997-06-06
 / EARLIER APPLICATION NUMBER: 60/068,368
 / EARLIER FILING DATE: 1997-12-19
 / NUMBER OF SEQ ID NOS: 118
 / SOFTWARE: PatentIn Ver. 2.0
 / SEQ ID NO 29
 / LENGTH: 879
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / US-09-152-060-29

Query Match 28.7%; Score 419.6; DB 4; Length 879;
 Best Local Similarity 78.2%; Pred. No. 4.9e-100;
 Matches 543; Conservative 0; Mismatches 144; Indels 7; Gaps 3;

QY 1 GGCCTGGCT 60
 DB 15 GGCCTGGCT 74
 QY 61 TGCCTGACTCAATGCT 120


```

Db      75 AGGGCTGACTCAGACCCCTCGGTGTCCAAAGGACTTGAAGACAGACCGCCACACTCAGCTG 134
Qy      121 CACTGGAACCCAGACGTCAGCTGGTGGTTATTAATGCTCTCTCTGTCACCAAGCACCC 180
Db      135 CACCGGGAACAACA--ACAAATGTGGCACCAGAGAGAGCTTGGCTGACAGACACCA 191
Qy      181 AGGCAAAAGCCCCCAACATCATTTATAGGTCAGTAATCGGCCCTCAGGGGGTTCTTAA 240
Db      192 GGGGCAACCTCCCAAACTCTCTCTCAAGAAATATAACCGGCCCTCAGGATCTCAGA 251
Qy      241 TCGCTTCTGTGCTCCAACTGTGGCAACAGGCTCTCCCTGACCATCTCTGGGCTTCAGGC 300
Db      252 GAGATTATCTGATCCAGGTCAGAGCCACATCTCTCCGACCATTAATCTGATCTCAAGCC 311
Qy      301 TGAGACGAGGCGATTAATTAATGCTGCTCATATCAAGAG--TACTTTCATGTCCT 357
Db      312 TGAGACGAGGCTGATTAATTAATGCTGCTCATATGACAGACCTTCGACGTTGGATGTT 371
Qy      358 CGGAACTGGGACCAAGGTCACCGCTCTAGTCAAGCCCAAGCCCAACCCCACTGTCACCT 417
Db      372 CGGCGAGGAGCAACAGCTACCGTCTAGTCAAGCCCAAGGCTGCGCCCTCGGTCACTCT 431
Qy      418 GTTCCCGCCCTCTCTGAGAGAGCTCCAAAGCCCAAGGCTCACTAGTGTCTGATCAG 477
Db      432 GTTCCCACTCTCTCTGAGAGAGCTTCAAGCCCAAGGCTCACTAGTGTCTGATCAG 491
Qy      478 TGACTTCTACCGGAGAGCTGTGACAGTGGCTGGAAGGAGATGAGCAGCCCGTCAAGGC 537
Db      492 TGACTTCTACCGGAGAGAGCTGTGACAGTGGCTGGAAGGAGATGAGCAGCCCGTCAAGGC 551
Qy      538 GGGAGTGAAGACCAACCACTCTCCAAAGAGCAACAAGTACGCGGCGAGCAGCTA 597
Db      552 GGGAGTGAAGACCAACCACTCTCCAAAGAGCAACAAGTACGCGGCGAGCAGCTA 611
Qy      598 CCGAGCCGGAAGCCGAGCAGTGAAGTCCCAAGAGCTACAGCTGCCAGTCAAGCA 657
Db      612 CCGAGCCGGAAGCCGAGCAGTGAAGTCCCAAGAGCTACAGCTGCCAGTCAAGCA 671
Qy      658 TGAAGGAGACCGGTGATGAAGACAGTGTGCC 691
Db      672 TGAAGGAGACCGGTGGA-GAAGACGTTGGCCCC 704

RESULT 13
US-09-152-060-47
; Sequence 47, Application US/09152060
; Patent No. 6448230
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P1.US
; CURRENT APPLICATION NUMBER: US/09/152,060
; EARLIER APPLICATION NUMBER: PCT/US98/04858
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/040,762
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/040,710
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/050,934
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,100
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,357
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,189
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/057,765
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/046,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/068,368
; EARLIER FILING DATE: 1997-12-19

```

```

; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 47
; LENGTH: 885
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-152-060-47

Query Match      28.6%; Score 418; DB 4; Length 885;
Best Local Similarity 78.1%; Pred. No. 1.3e-99;
Matches 542; Conservative 0; Mismatches 145; Indels 7; Gaps 3;

Qy      1 GGGCTGGGCTCTGCTATTCCTCAACCTCTCTCACTCAGGACAGAGGCTCTGGGCCAGTC 60
Db      23 GCGCTGGGCTCTGCTCTCTCTCTGACCTCTCTCACTCAGTGTGAGTGTCCAGGC 82
Qy      61 TGCCCTGACTCACTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db      83 AGGGCTGACTCAACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 142
Qy      121 CACTGGAACCCAGACGTCAGCTGGTGGTTATTAATGCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Db      143 CACCGGGAACAACA--ACAAATGTGGCACCAGAGAGAGCTTGGCTGACAGACCA 199
Qy      181 AGGCAAAAGCCCCCAACATCATTTATAGGTCAGTAATCGGCCCTCAGGGGTTCTTAA 240
Db      200 GGGCCACCTCTCCAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 259
Qy      241 TCGCTTCTGTGCTCCAACTGTGGCAACAGGCTCTCCCTGACCATCTCTGGGCTTCAGGC 300
Db      260 GAGATTATCTGACATCAAGTTCAGAGCCACATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 319
Qy      301 TGAGACGAGGCTGATTAATTAATGCTGCTCATATCAAGAG--TACTTTCATGTCCT 357
Db      320 TGAGACGAGGCTGATTAATTAATGCTGCTCATATGACAGAGCTGCGATTTGATGTT 379
Qy      358 GGGAACTGGGACCAAGTCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 417
Db      380 CGGCGAGGAGCAACCACTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 439
Qy      418 GTTCCCGCCCTCTCTGAGAGGCTCCAGGCAACAAGGCAACAAGTGTGTCTGATCAG 477
Db      440 GTTCCCACTCTCTCTGAGAGGTTCAAGCCCAACAAGGCTCACTGTGTCTAATAG 499
Qy      478 TGACTTCTACCGGAGAGCTGTGACAGTGGCTTGAAGGCAAGATGCAAGCCCGTCAAGGC 537
Db      500 TGACTTCTACCGGAGAGCGGTGACAGTGGCTTGAAGGCAAGATGCAAGCCCGTCAAGGC 559
Qy      538 GGGAGTGAAGACCAACCACTCTCCAAAGAGCAACAAGTACGCGGCGAGCAGCTA 597
Db      560 GGGAGTGAAGACCAACCACTCTCCAAAGAGCAACAAGTACGCGGCGAGCAGCTA 619
Qy      598 CCGAGCTTGAAGCCCGAGAGTGAAGTCCCAAGAGCTACAGCTGCGAGTCAAGCA 657
Db      620 CCGAGCTTGAAGCCCGAGAGTGAAGTCCCAAGAGCTACAGCTGCGAGTCAAGCA 679
Qy      658 TGAAGGAGACCGGTGATGAAGACAGTGTGCC 691
Db      680 TGAAGGAGACCGGTGGA-GAAGACGTTGGCCCC 712

RESULT 14
US-08-523-894-5
; Sequence 5, Application US/08523894
; Patent No. 6136310
; GENERAL INFORMATION:
; APPLICANT: Hanna, Nabli
; APPLICANT: Newman, Roland A.
; APPLICANT: Reff, Mitchell E.
; TITLE OF INVENTION: Recombinant Anti-CD4 Antibodies for Human
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:

```

```

ADDRESSER: BURNS, DOANE, SWECKER & MATHIS
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314-3187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/523,894
FILING DATE: 06-SEP-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 012712-165
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ. ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 702 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
POSITION IN GENOME:
CHROMOSOME/SEGMENT: lambda variable and constant domains in
FEATURE:
NAME/KEY: CDS
LOCATION: 1..702
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..702
US-08-523-894-5

Query Match      28.4%; Score 414; DB 3; Length 702;
Best Local Similarity 78.0%; Pred. No. 1.3e-98;
Matches 541; Conservative 0; Mismatches 140; Indels 13; Gaps 3;

QY 1 GGGCTGGGCTCTGCTATTCCTCAACCTCTCACTAGGACAGAGGCTCTGGGCCAGTC 60
DB 3 GGCTGGGCTCTGCTCTGCTCTCGGCTCTGCTCACTTACAGACTGCGGCTCTCTA 62
QY 61 TGGCCCTGACTAGTCTGCTCCGCTGCTGGTCTCTGAGACGTGATACATCTCTG 120
DB 63 TGAGTTGAGTCACTGCTGCTCACTGCTCGTGTCCAGAGACAGCGCGGTTCACTG 122
QY 121 CACTGAAACAGACAGTCACTGCTGTTAATACTATGCTCTCTGTACCAACAGACCC 180
DB 123 TGGGGGAGAAAC-----GTGGAAGAAAAGTGAAGTGTACCAAGAGCC 173
QY 181 AGGCAAGCCCCCAAACTATCATTTATGAGTCACTAATCGGCTCTCAGGGTTTCTA 240
DB 174 ACCGCAAGCCCTGCTGTGCTATCTAATGTCAGACGAGCGCCCTCAGGATCCCTGC 233
QY 241 TCGCTTCTCTGCTGCTCAAGTCTGGAACACGCGCTCTGACCATCTCTGGGCTCCAGG 300
DB 234 GCGATTCTCTGCTCTCACTCAAGGAAACCGCCCACTGACCATAGGGGGTGAAGGC 293
QY 301 TGAAGACGAGGCTGATTATTACTGCTGCTCATATTAACAAGACTTCTCAT---GTCT 357
DB 294 CCGGGATGAGGCTGACTATTACTGTCAAGTGTGGAACAGTACTGATGATTTGGTCTT 353
QY 358 CGGAATCGGAGCAAGGTCACGCTCTAGGTCAGCCCAAGGCAAGCCCACTGTCACCT 417
DB 354 CCGGAGAGGAGCCCGGCTGACCTCTTAGTTCAGCCCAAGGCTGCCCTCTCGTCACTCT 413

```

```

QY 418 GTTCCGCGCTCTCTGAGAGAGCTCCAGCCAAAGGCGACACTAGTGTGATCAG 477
DB 414 GTTCCGCGCTCTCTGAGAGAGCTTCAAGCCAAAGGCGCACCTGATGTCTTAAAG 473
QY 478 TGACTTCACTCCGAGAGCTGTGACAGTGGCTTGAAGGACAGATGAGAGCCCCGTAAGGC 537
DB 474 TGACTTCACTCCGAGAGCGGTGACAGTGGCTTGAAGGACAGATGAGAGCCCCGTAAGGC 533
QY 538 GGAAGTGAAGACCAACCAACCTTCCAAAGAGACCAACAAATAGCGGCGACAGCTA 597
DB 534 GGAAGTGAAGACCAACCAACCTTCCAAAGAGACCAACAAATAGCGGCGACAGCTA 593
QY 598 CTGAGCCTGACCGCCGACGAGTGAAGTCCACAGAACTACAGCTGCCAGGTACAGCA 657
DB 594 CTGAGCCTGACCGCTGAGAGAGTGAAGTCCACAGAACTACAGCTGCCAGGTACAGCA 653
QY 658 TGAAGGAGCACCGTGTGATGAAGACAGTGTCC 691
DB 654 TGAAGGAGCACCGTGTGA-GAAGACAGTGTGCC 686

```

```

RESULT 15
US-08-487-550-1
Sequence 1, Application US/08487550
Patent No. 611898
GENERAL INFORMATION:
APPLICANT: Anderson, Darrell R.
TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC
TITLE OF INVENTION: TO HUMAN B7.1 AND/OR B7.2 PRIMATED FORMS THEREOF.
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS
NUMBER OF INVENTION: IMMUNOSUPPRESSANTS"
CORRESPONDENCE ADDRESSES:
ADDRESSER: BURNS, DOANE, SWECKER & MATHIS
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,550
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 012712-131
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-6620
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 705 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
MOLECULE TYPE: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..705
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..705
US-08-487-550-1

Query Match      28.1%; Score 410.8; DB 3; Length 705;
Best Local Similarity 77.7%; Pred. No. 9e-98;

```

Matches 539; Conservative 0; Mismatches 142; Indels 13; Gaps 3;

```
QY      1 GGCCTGGGCTCTGCTATTCTCTCACTCTCACTGAGGACAGAGTCTCTGGGCCAGTC 60
Db      6 GGTCCCCGCTCAGCTCTGAGGCTCTGCTCTGCTCTGCTCCAGGTGACAGATGACCTA 65
QY     61 TGGCTGACTCAGTCTGCTCCGTCGTCTGCTCTGCTCTGCTGAGTCACTCATCTCTG 120
Db     66 TGAATGACTCAGTCTGCTCCGTCGTCTGCTCTGCTCTGCTGAGTCACTCATCTCTG 125
QY     121 CACTGGAACAGCAGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 180
Db     126 TGGGGAGAGACAGTCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 176
QY     181 AGGCAAGAGGCTCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 240
Db     177 AGGCGGGGCTCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 236
QY     241 TCGCTTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 300
Db     237 GCGATTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 296
QY     301 TGAAGACAGAGTCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 357
Db     297 CGGGGATGAGGCTCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 356
QY     358 CGGAATGGAACAGTCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 417
Db     357 CGAGAGAGGAGTCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 416
QY     418 GTTCCCGCTCTCTGAGAGTCTCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCTG 477
Db     417 GTTCCCGCTCTCTGAGAGTCTCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCTG 476
QY     478 TGAATCTCAAGTCTGAGAGTCTCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCT 537
Db     477 TGAATCTCAAGTCTGAGAGTCTCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCT 536
QY     538 GGAAGTGAAGACCAAGTCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 597
Db     537 GGAAGTGAAGACCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 596
QY     598 CTTGAGCTGAGTCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 657
Db     597 CTTGAGCTGAGTCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 656
QY     658 TGAAGGAGCAGTCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 691
Db     657 TGAAGGAGCAGTCAAGTCAAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 689
```

Search completed: April 26, 2004, 20:28:40
Job time: 124.019 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 15:46:02 ; Search time 3550.63 Seconds

(Without alignments)
12279.158 Million cell updates/sec

Title: US-10-076-747-8

Perfect score: 1460

Sequence: 1 ggcctggagcctgctatcc.....ggcagcagagccacgcagct 1460

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vtc:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	638.8	43.8	889	12	BG756342
2	637.6	43.3	786	12	BG758901
3	637.8	43.3	911	13	BQ712542
4	617.6	42.3	829	12	BG685732

5	615.6	42.2	868	12	BG483745
6	614.8	42.1	810	12	BG754419
7	613	42.0	729	12	BF126874
8	612.4	41.9	808	12	EM007694
9	610.8	41.8	1049	12	EM920304
10	607.8	41.6	686	12	BG759257
11	606.8	41.6	957	13	BX446259
12	606.8	41.6	1150	13	BX449367
13	606.4	41.5	1201	13	BX462166
14	604	41.4	1201	13	BX442894
15	603.8	41.4	840	12	BG397282
16	602.8	41.3	903	13	BE561081
17	602	41.2	945	13	BX446252
18	601.6	41.0	1040	12	BX410975
19	599.2	41.0	1063	12	EM920655
20	598.8	41.0	739	10	BE397699
21	598.2	41.0	813	10	BF974961
22	598.2	41.0	953	10	BF976229
23	597.6	40.9	958	13	BX426264
24	597	40.9	1019	10	BE271225
25	596.4	40.8	767	9	AV697043
26	595.4	40.8	884	12	EM051870
27	595.2	40.8	892	12	B1227090
28	595	40.8	787	12	B1820758
29	594.8	40.7	898	13	BQ711473
30	593.2	40.6	869	13	BE562370
31	593.2	40.6	950	12	EM918781
32	593	40.6	851	10	BE559866
33	591.4	40.5	850	10	BF975248
34	591	40.5	854	12	B1227345
35	590.6	40.5	876	12	EM051654
36	590.2	40.4	772	10	BE397742
37	589.8	40.4	786	12	B1224785
38	589.2	40.4	918	12	B1823305
39	589	40.3	749	12	B1226887
40	589	40.3	908	12	BG756874
41	588	40.3	737	14	CD691527
42	587.8	40.3	880	10	BF975970
43	587.4	40.2	819	10	BE560212
44	587.2	40.2	931	12	EM051920
45	586.4	40.2	856	10	BE561313

ALIGNMENTS

RESULT 1
BG756342
LOCUS 602713662P1 NIH_MGC_48 Homo sapiens CDNA clone IMAGE:4853919 5',
DEFINITION
ACCESSION BG756342
VERSION BG756342.1 GI:14066995
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 889)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rudin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lln.gov
Plate: LNCMI700 row: 1 column: 16
High quality sequence stop: 810.

DB 334 TGAGGATGAGCTGATTAATTACTGACGCTCATATGACAGGACGACCAATTAATGCTTCGG 393
QY 361 AACTGGAGACCAAGTCAACCGCTAGTGTACAGCCCAAGGCGCAACCCCACTGACCTGTT 420
DB 394 AACTGGAGACCAAGTCAACCGCTAGTGTACAGCCCAAGGCGCAACCCCACTGACCTGTT 453
QY 421 CCCGCTCTCTGAGAGCTCCAAAGCCCAAGGCGCAACCTGATGTGTATGATGATGA 480
DB 454 CCCGCTCTCTGAGAGCTCCAAAGCCCAAGGCGCAACCTGATGTGTATGATGATGA 513
QY 481 CTTCTAACCGGAGAGCTGTACAGTGGCTTGGAGGAGCAAGTGGACGCCCCGTCAGGCGGG 540
DB 514 CTTCTAACCGGAGAGCTGTACAGTGGCTTGGAGGAGCAAGTGGACGCCCCGTCAGGCGGG 573
QY 541 AGTGAAGACCAACCAACCTCCAAACAGAGCAACAAGTACGCGGCGCAGAGCTACT 600
DB 574 AGTGAAGACCAACCAACCTCCAAACAGAGCAACAAGTACGCGGCGCAGAGCTACT 633
QY 601 GAGCTGACGCGCCGAGAGCTGGAAGTCCCAAGAACTTACAGCTGCCAGTTCACGATGA 660
DB 634 GAGCTGACGCGCCGAGAGCTGGAAGTCCCAAGAACTTACAGCTGCCAGTTCACGATGA 693
QY 661 AGGAGACACCGTGGATGAAGACAGTGTGCC 691
DB 694 CGGAGACCGTGGAG-GAAGACAGTGTGCC 723

RESULT 3
LOCUS B0712542 911 bp mRNA linear EST 16-JUL-2002
DEFINITION AGENCOURT_8495253 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6302004
ACCESSION B0712542
VERSION B0712542
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 911)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Mark Watson
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L10M2518 row: m column: 13
High quality sequence stop: 650.
Location/Qualifiers
1. 911
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6302004"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_113"
/note="Organ: spleen; Vector: pOT7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGGACAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN
Query Match 43.3%; Score 631.8; DB 13; Length 911;

Best Local Similarity 96.0%; Pred. No. 8.5e-134;
Matches 648; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 1 GGCCTGGCTCTGCTATTCCTGACCCCTGCTGCTGAGGACAGAGGATCCCTGGCCAGTGC 60
DB 37 GGCCTGGCTCTGCTATTCCTGACCCCTGCTGCTGAGGACAGAGGATCCCTGGCCAGTGC 96
QY 61 TGCCTGACTCAGTCTGCTCCGCTGTCTGGGTTCTCTGGAACAGTCAGTCACTCTCTG 120
DB 97 TGCCTGACTCAGCTGCTGCTCCGCTGTCTGGGTTCTCTGGAACAGTCAGTCACTCTCTG 156
QY 121 CACTGGAACCAAGATGACGTTGGTGTATTAATTAATTAATTAATTAATTAATTAATTA 180
DB 157 CAGTGAACCAAGATGACGTTGGTGTATTAATTAATTAATTAATTAATTAATTAATTA 216
QY 181 AGGCAAGCCCCCAATCATATTATAGAGTCAGTAAATCGGCTCTGAGGGTTTCTAA 240
DB 217 AGGCAAGCCCCCAATCATATTATTTATATATATATATATATATATATATATATATAT 276
QY 241 TCGCTTCTGCTGCTCAAGTGTGGCAACAGGCTCTGACCATCTCTGGGCTCAAGG 300
DB 277 TCGCTTCTGCTGCTCAAGTGTGGCAACAGGCTCTGACCATCTCTGGGCTCAAGG 336
QY 301 TGAAGACAGGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 360
DB 337 TGAAGACAGGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 396
QY 361 AACTGGAGACCAAGTCAACCGCTAGTCAAGCCCAAGGCGCAACCCCACTGCTGTT 420
DB 397 AACTGGAGACCAAGTCAACCGCTAGTCAAGCCCAAGGCGCAACCCCACTGCTGTT 456
QY 421 CCCGCTCTCTGAGAGGCTTCCAAAGCCCAAGGCGCAACCTGATGTGTATGATGA 480
DB 457 CCCGCTCTCTGAGAGGCTTCCAAAGCCCAAGGCGCAACCTGATGTGTATGATGA 516
QY 481 CTTCTAACCGGAGAGCTGTGACAGTGGCTTGGAGGAGATGAGGCGCCGTCAGGCGGG 540
DB 517 CTTCTAACCGGAGAGCTGTGACAGTGGCTTGGAGGAGATGAGGCGCCGTCAGGCGGG 576
QY 541 AGTGAAGACCAACCAACCTCCAAACAGAGCAACAAGTACGCGGCGCAGCTACT 600
DB 577 AGTGAAGACCAACCAACCTCCAAACAGAGCAACAAGTACGCGGCGCAGCTACT 636
QY 601 GAGCTTACGCGCCGAGAGTGAAGTCCCAAGAGTACAGTGCAGGTCAGGATGA 660
DB 637 GAGCTTACGCGCCGAGAGTGAAGTCCCAAGAGTACAGTGCAGGTCAGGATGA 696
QY 661 AGGAGACACCGTGGAG 675
DB 697 AGGAGACACCGTGGAG 711

RESULT 4
LOCUS B6685732 829 bp mRNA linear EST 01-MAY-2001
DEFINITION 602637827F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4765447 5',
mRNA sequence.
ACCESSION B6685732
VERSION B6685732
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 829)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Louis M. Straudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM1624 row: f column: 08
 High quality sequence stop: 821.
 Location/Qualifiers

FEATURES

source

1..829
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4765447"
 /issue_type="primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH_MGC_46"
 /note="Organ: B-cells; Vector: pOTB1; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dt priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."

ORIGIN

Query Match

Best Local Similarity 94.2%; Score 617.6; DB 12; Length 829;

Matches 652; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY 1 GGCTTGGGCTCTGCTATTCCTCAACCTCTCACTCAAGGACAGGGTCTGGGCCCAAGTC 60
 DB 34 GGCTTGGGCTCTGCTCTCTCAACCTCTCACTCAAGGACAGGGTCTGGGCCCAAGTC 93
 QY 61 TGGCTTGGGCTCTGCTCTCTCAACCTCTCACTCAAGGACAGGGTCTGGGCCCAAGTC 120
 DB 94 TGGCTTGGGCTCTGCTCTCTCAACCTCTCACTCAAGGACAGGGTCTGGGCCCAAGTC 153
 QY 121 CACTGGAACGAGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 DB 154 CACTGGAACGAGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 213
 QY 181 AGGCAAGCCGCCCAAACTCACTTTATGAGGTCATTAATCGGCCCTCAGGGGTTCTTA 240
 DB 214 AGGCAAGCCGCCCAAACTCACTTTATGAGGTCATTAATCGGCCCTCAGGGGTTCTTA 273
 QY 241 TCGCTTCTGCTGCTCAAGTCTGCAACAGGCTCTGACCATCTGAGGCTCCAGGC 300
 DB 274 TCGCTTCTGCTGCTCAAGTCTGCAACAGGCTCTGACCATCTGAGGCTCCAGGC 333
 QY 301 TGAAGAGAGAGCTGATTATTACTGCTGCTATTAACAAGATCTTCAATGCTCTTGG 360
 DB 334 TGAAGAGAGAGCTGATTATTACTGCTGCTATTAACAAGATCTTCAATGCTCTTGG 393
 QY 361 AACTGGAGCAAGGTCACTCTAGGTCAAGCCCAAGGCCCAACCCCACTGCTCTGTT 420
 DB 394 AACTGGAGCAAGGTCACTCTAGGTCAAGCCCAAGGCCCAACCCCACTGCTCTGTT 453
 QY 421 CCGGCTCTCTTGAAGAGTCTCAAGCCCAAGGCCCAACCTGATGCTGATCACTGA 480
 DB 454 CCGGCTCTCTTGAAGAGTCTCAAGCCCAAGGCCCAACCTGATGCTGATCACTGA 513
 QY 481 CTCTTACCGGAGAGCTGTGACATGAGGCTGGAAGGAGATGAGGAGCCCGTCAAGGCGGG 540
 DB 514 CTCTTACCGGAGAGCTGTGACATGAGGCTGGAAGGAGATGAGGAGCCCGTCAAGGCGGG 573
 QY 541 AGT-GGAGACCAACCAACCTCCAAACAGAGCAACCAAGTACGGGAGCAGAGCTACC 599
 DB 574 AGTGGAGACCAACCAACCTCCAAACAGAGCAACCAAGTACGGGAGCAGAGCTACC 633
 QY 600 TGAAGCTGAGCCGAGAGTGAAGTCCAGAGGCTACAGTCCCAAGTCCAGGCTCAAGCAG 659
 DB 634 TGAAGCTGAGCCGAGAGTGAAGTCCAGAGGCTACAGTCCCAAGTCCAGGCTCAAGCAG 693

QY 660 AAGGAGACCCGTCGATGAAGACAGTGTCTCC 691
 DB 694 AAGGAGACCCGTCGATGAAGACAGTGTCTCC 725

FEATURES

source

RESULT 5
 LOCUS BG483745 868 bp mRNA linear EST 21-MAR-2001
 DEFINITION 60250383F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4616829 5',
 mRNA sequence.
 ACCESSION BG483745
 VERSION BG483745.1 GI:13415934
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 868)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: CLONTECH Laboratories, Inc.
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM1372 row: e column: 22
 High quality sequence stop: 834.
 Location/Qualifiers

ORIGIN

source

1..868
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4616829"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_1lb="NIH_MGC_77"
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:
 SfiI (ggccgctggcc); Site 2: SfiI (ggccatgagcc); 5' and
 3' adaptors were used in cloning as follows: 5' adaptor
 sequence: 5'-CAAGGCAATATGAGC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCCGAGGCGGCGGACATG-dt(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.9
 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC library."

ORIGIN

Query Match

Best Local Similarity 92.7%; Score 615.6; DB 12; Length 868;

Matches 669; Conservative 0; Mismatches 49; Indels 4; Gaps 2;

QY 1 GGCTTGGGCTCTGCTATTCCTCAACCTCTCACTCAAGGACAGGGTCTGGGCCCAAGTC 60
 DB 42 GGCTTGGGCTCTGCTATTCCTCAACCTCTCACTCAAGGACAGGGTCTGGGCCCAAGTC 101
 QY 61 TGGCTTGGGCTCTGCTCTCTCAACCTCTCACTCAAGGACAGGGTCTGGGCCCAAGTC 120
 DB 102 TGGCTTGGGCTCTGCTCTCTCAACCTCTCACTCAAGGACAGGGTCTGGGCCCAAGTC 161
 QY 121 CACTGGAACGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 DB 162 CACTGGAACGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 221
 QY 181 AGGCAAGCCGCCCAAACTCACTTTATGAGGTCATTAATCGGCCCTCAGGGGTTCTTA 240
 DB 222 AGGCAAGCCGCCCAAACTCACTTTATGAGGTCATTAATCGGCCCTCAGGGGTTCTTA 281

QY	241	TCGCTTCTCTGGCTCCAGGCTTGCGCAACGGCCCTCCGTGACCACTCTGGGGCTCCAGG	300
Db	282	TCCGCTCGCTGGCTCCAGTCTGGGAAACAGGCCCTCCGTGACCACTCTGGGGCTCCAGG	341
QY	301	TGAAGACGAGGCTGATTATTACTGCTGCTCATATACAAAGATCTTTCATGTCTTCGG	360
Db	342	TGAGGACGAGGCTGATTATTACTGCGAGCTCATACAAAGACGACCACTCGAGTCTTCGG	401
QY	361	AACTGGAGCAAGGTCAACGTCCTTAGTCAAGCCCAAGGCCCAACCCCACTGTCACTGTGT	420
Db	402	AACTGGAGCAAGGTCAACGTCCTTAGTCAAGCCCAAGGCCCAACCCCACTGTCACTGTGT	461
QY	421	CCCGCCCTCTCTGTGAGAGCTCCAAAGCCAAACAAGCCACACTAGTGTGTCTGATCACTGA	480
Db	462	CCCGCCCTCTCTGTGAGAGCTCCAAAGCCAAACAAGCCACACTAGTGTGTCTGATCACTGA	521
QY	481	CTTTACCCGGAGGCTGTGACAGTGGCTGGAAAGGCAATGGCAGCCCGTCAAGGACGG	540
Db	522	CTTTACCCGGAGGCTGTGACAGTGGCTGGAAAGGCAATGGCAGCCCGTCAAGGACGG	581
QY	541	AGTGGAGACCAACCAACCTCTCCAAACAGAGCAACCAAGTACGGGGCCAGCAAGTACTT	600
Db	582	AGTGGAGACCAACCAACCTCTCCAAACAGAGCAACCAAGTACGGGGCCAGCAAGTACTT	641
QY	601	GAGCCTGACGCCCGAGCA--GTGGAAGTCCCAACAAGTACAGTGCACAGTCAAGC--	656
Db	642	GAGCCTGACGCCCGAGCAAGTGAAGATCCCAACAAGTACAGTGCACAGTCAAGC--	701
QY	657	ATGAAGGAGACCACTGATGATGAAGACAGTGTCCCTCCACAGAAATGTTCTATAGGTTTC	716
Db	702	TTGAAGGAGACCACTGATGATGAAGACCAAGTGGCCCTTACAGATGTCAATTAAGTTCC	761
QY	717	TA 718	
Db	762	AA 763	

RESULT 6
BG754419
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

810 bp mRNA linear EST 15-MAY-2001
602710036r1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4846521 5',
mRNA sequence.
BG754419
BG754419.1 GI:14065072
EST.
Homo sapiens (human)
Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eumetazoa; Eutheria; Primates; Carnivora; Hominidae; Homo.
1 (bases 1 to 810)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: CGAPbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L16C1687 row: h column: 10
High quality sequence scop: 805.

```

FEATURES
    source
        Location/Qualifiers
            1..810
            /organism="Homo sapiens"
            /mol_type="RNA"
            /db_xref="taxon:9606"
            /db_xref="IMAGE:4846521"
            /clone_type="Primary B-cells from tonsils (cell line)"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NH1_WGC_48"

```

```

/Note:"Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;
Site:2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

Query Match	Similarity	42.1%;	Score 614.8;	DB 12;	Length 810;
Best Local	Similarity	42.1%;	Score 614.8;	DB 12;	Length 810;
Matches	663;	Conservative	0;	Mismatches	22;
				Indels	5;
				Gaps	3;
QY	1	GGCTTGAGGCTCTGCTATTTCTCAACCTCTCTCACTCAGGGGACAGAGGCTCTGGGCCCCAGTC	60		
Db	18	GGCTTGAGGCTCTGCTGCTCTCTCACTCTCTCACTCAGAGACA-GGGCTTGAGGCCCCAGTC	76		
QY	61	TGACCTGACTCAGTCTCCTCCGNGTCTGGGCTCTCCCTGACAGATCGATACCACTCTCTG	120		
Db	77	TGCCCTGACTCAGCTGCTCCTCGGTGTGGGTCTCTCGACAGATCGATACCACTCTCTG	136		
QY	121	CAGTGAACAGCAAGTCAAGTTGGTGTAACTATGTCTCTGGTACCAAGACACC	180		
Db	137	CAGTGAACAGCAAGTATGTTGGAGTTAACTTGTCTCTGGTACCAAGACACC	196		
QY	181	AGGCAAGAGCCCCCAAACTCATATTAATGAGTACAGTAATCGAGCCCTCAGAGGTTTCTAA	240		
Db	197	AGGCAAGAGCCCCCAAACTCATATTAATGAGGCAAGTAAGGAGCCCTCAGAGGTTTCTAA	256		
QY	241	TCGGTTCTTGCTCTCCAAAGTCTGGCAACAAGGCTCTCTGACCATCTCTGGGCTCCAGGC	300		
Db	257	TCGGTTCTTGCTCTCCAAAGTCTGGCAACAAGGCTCTCTGACCAATCTCTGGGCTCCAGGC	316		
QY	301	TGAGAGAGAGGCTATTTACTAGCTGCTCATATPACAAGAAGTACTCTATGCTCTCGG	360		
Db	317	TGAGAGAGAGGCTATTTACTAGCTGCTCATATGCAAGTACGCTTT--ATGCTTCCG	373		
QY	361	AAGTGGAGCAAGGTCAACGCTCTAGGCTCAGCCCAAGGCCCAAGCCCACTGTCACTCTGTT	420		
Db	374	AAGTGGAGCAAGGTCAACGCTCTAGGCTCAGCCCAAGGCCCAAGCCCACTGTCACTCTGTT	433		
QY	421	CCGGGCTCTCTGAGAGGCTCCAAAGCAACAAGGCAACACTATGCTGTCTGATCAATGA	480		
Db	434	CCGGGCTCTCTGAGAGGCTCCAAAGCAACAAGGCAACACTATGCTGTCTGATCAATGA	493		
QY	481	CTTCTAACCCGGAGACTGTGACAGTGGCTGGAAAGGCAGATGAGCAAGCCCGTCAAGGCGGG	540		
Db	494	CTTCTAACCCGGAGACTGTGACAGTGGCTGGAAAGGCAGATGAGCAAGCCCGTCAAGGCGGG	553		
QY	541	AGTGAACAACAACCAACCTCCAAACAGAGCAACAACAATACGGGGCCAGAGCTACCT	600		
Db	554	AGTGAACAACAACCAACCTCCAAACAGAGCAACAACAATACGGGGCCAGAGCTACCT	613		
QY	601	GAGCCTGACGCCCGAGAGTGAAGTCCACAGAACTACAGTGTCCAGGTCAAGCATGA	660		
Db	614	GAGCCTGACGCCCGAGAGTGAAGTCCACAGAACTACAGTGTCCAGGTCAAGCATGA	673		
QY	661	AGGAGAGCAAGCTGATGAAGAGCAAGTGTCC	690		
Db	674	AGGAGAGCAAGCTGGA-GAAGCAAGTGTCC	702		

RESULT	7
BF128674	
LOCUS	729 bp mRNA linear EST 24-OCT-2000
DEFINITION	60181093FF1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4054065 5',
ACCESSION	BF128674
VERSION	BF128674.1 GI:10967714
KEYWORDS	EST.
SOURCE	Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 729)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC) JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rs9ab@biml.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
Plate: LINC893 row: m column: 10
High quality sequence stop: 696.
Location/Qualifiers
1..729
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4054065"
/tissue_type="Primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_id="NH MGC 48"
/note="Organ: B-cells; Vector: pONB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GAGCAGG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NH_MGC Library."

ORIGIN

Query Match	42.0%;	Score 613;	DB 10;	Length 729;
Best Local Similarity	96.7%;	Pred. No. 1.6e-129;		
Matches 670;	Conservative 0;	Mismatches 15;	Indels 8;	Gaps 4;
QY 1	GGCCTGGGCTCGCATTCCTGACCCCTCTCACTGAGGCAAGGGTCTGGGCCAGTC	60		
DB 1	GGCCTGGGCTCGCATTCCTGACCCCTCTCACTGAGGCAAGGGTCTGGGCCAGTC	60		
QY 61	TGCCCTGACTAGTGCCTCCGTCTGGGCTCTCTGGACAGTGATCAACATCTCTG	120		
DB 61	TGCCCTGACTAGTGCCTCCGTCTGGGCTCTCTGGACAGTGATCAACATCTCTG	120		
QY 121	CACCTGGAACGACATCACTGATGAGTTAACTATCTCCCTGACCAAGCAGACC	180		
DB 121	CACCTGGAACGACATCACTGATGAGTTAACTATCTCCCTGACCAAGCAGACC	180		
QY 181	AGGCAAGCCCAACTCATATTGAGGTGAGTATGAGCCCTCAGGAGTTCTTAA	240		
DB 181	AGGCAAGCCCAACTCATATTGAGGTGAGTATGAGCCCTCAGGAGTTCTTAA	240		
QY 241	TGCTTTCTTGCTCCAAAGTGTGGCAACGGCTTCCTGACCATCTTGAGCTCAGGC	300		
DB 241	TGCTTTCTTGCTCCAAAGTGTGGCAACGGCTTCCTGACCATCTTGAGCTCAGGC	300		
QY 301	TGAGACAGAGGTGATTATCTGCTGCATATACA--AGAGTACTTCTCATGTCTT	357		
DB 301	TGAGACAGAGGTGATTATCTGCTGCATATACA--AGAGTACTTCTCATGTCTT	357		
QY 361	CGGAAGTGGACCAAGGTCACTCTCTGATGATGAGCCCAAGGCAACCCCATCTCACT	420		
DB 361	CGGAAGTGGACCAAGGTCACTCTCTGATGATGAGCCCAAGGCAACCCCATCTCACT	420		
QY 418	GTTCGCGCCCTCTCTGAGAGCTCCAAAGCCAAAGGCACTAGTGTCTGATAG	477		
DB 418	GTTCGCGCCCTCTCTGAGAGCTCCAAAGCCAAAGGCACTAGTGTCTGATAG	477		
QY 480	GTTCGCGCCCTCTCTGAGAGCTCCAAAGCCAAAGGCACTAGTGTCTGATAG	480		
DB 480	GTTCGCGCCCTCTCTGAGAGCTCCAAAGCCAAAGGCACTAGTGTCTGATAG	480		

Oy		478	TGACCTTACCCCGAGACTTGACAGGGCCTGAAAGGCAGATGGCAGCCCCGTCAAGGC	537
Db		481	TGACTTTACCCGGAGCTGTGACAGTGGCTGGAAGGACGATGGCAGCCCCGTCAAGGC	540
Oy		538	GGAAGTGAACAACCAACCCCTTCMAACAGACACACAAAGTAAGCGGCCGACGACTA	597
Db		541	GGAAGTGAAG---ACACAAACCCTCCAAAAGAACCAACAAAGTACCGCG--CAGCAGCTA	596
Oy		598	CCTGAGCCTGACGCCCGACAGAGTGAAGTCCCAAGAGCTACAGCTGCCAGGTCAAGCA	657
Db		597	CTTAGACCTAGCGCCGACAGAGTGAAGT-CGACAGAGCTACAGCTGCCAGGTCAAGCA	655
Oy		658	TGAAGGAGCACCGTGGATGAAGACAGTGTCC	690
Db		656	TGAAGGAGCACCGTGGACAGAGAAGTGGCCC	688
RESULT 8				
LOCUS	BMO07694		808 bp	mRNA linear EST 30-OCT-2001
DEFINITION	603617126F1 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:5440855 5',			
ACCESSION	BMO07694			
KEYWORDS	mRNA sequence.			
SOURCE	BMO07694.1 GI:16522048			
ORGANISM	Homo sapiens (human)			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
JOURNAL	1 (bases 1 to 808)			
COMMENT	NIH-MGC http://mgs.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: gga@ds-remail.nih.gov Tissue Procurement: Dr. Mark Watson cDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov Plate: LINC914 row: d column: 08 High quality sequence stop: 806. Location/Qualifiers			
FEATURES				
source				
	1..808			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:5440855"			
	/lab_host="DH10B (phage-resistant)"			
	/clone_lib="NIH_MGC_113"			
	/note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAACGAC(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: This is a NIH_MGC library."			
ORIGIN				
Query Match				
Best Local Similarity	91.9%; Score 612.4; DB 12; Length 808;			
Matches 648; Conservative 0; Mismatches 41; Indels 1; Gaps 1;				
Oy	1	GAGCTGGGCTCTGCTATTCTCTCAACCTCTCTCACTCAGGACACAGGATCTGGGGCCCATGC	60	
Db	33	GAGCTGGGCTCTGCTCTCTCTCTCAAGCCTCTCACTCAAGCCACAGATCCCTGGGCTCAGTC	92	
Oy	61	TGCCCTACCTCAGCTGCGGCTCGGAGTGTGGGCTCTCCCGACAGCATCAACCATCTCTCTG	120	

Db	93	TGCGCTGACGACAGCCCTCGGTGAGTGTCCGGGGTCTCTGAGACAGTCAAGTACATCTCTCG	152
Qy	121	CATGTGAACCGACGAGTCAAGCTTGGTGGTTATTAATAATGTCTCTGTGGTACCAAGCAAGCC	180
Db	153	CAGTGAACCAACAGAGTGGTGGATTAAACATCTGTCTCTGTGGTACCAAGCAAGCC	212
Qy	181	AGCAAGAGCCCAAACTCATTTATGAGTCAAGTAATGGGCGCTCAGGGGTTTCTAA	240
Db	213	AGCAAGAGCCCAAACTCATTTATGAGTCAAGTAATGGGCGCTCAGGGGTTTCTTA	272
Qy	241	TGCGTCTCTGTGGTCCAAAGTCTTGGCAACAGGCGCTCCCTGACATCTTGGGGTCCAGGC	300
Db	273	TCCCTTCTGTGGTCCAAAGTCTTGGCAACAGGCGCTCCCTGACATCTTGGGGTCCAGGC	332
Qy	301	TGAGGACGAGGGCTGATTATTAATGCTGTGCATATACAAAGAAAGTACTTTCATGTCTTGG	360
Db	333	TGAGGATGAGGCTGATTATTAATGCTGTGCATATACAAAGAAAGTACTTTCATGTCTTGG	392
Qy	361	AACTGGAGCAAGATCAAGTCCAGTCTTAGATCAAGCCCAAGGCCAACCCCACTGTCACTGT	420
Db	393	AACTGGAGCAAGTCAAGTCCAGTCTTAGATCAAGCCCAAGGCCAACCCCACTGTCACTGT	452
Qy	421	CCGGCCCTCTCTGAGGAGTCTCCAGGCCAACGAGGCCACATAGTGTGTGATCAAGTAA	480
Db	453	CCGGCCCTCTCTGAGGAGTCTCCAGGCCAACGAGGCCACATAGTGTGTGATCAAGTAA	512
Qy	481	CTTCTTCCCGGAGCTGTGACAGTGGCTTGGAGGAGCATGTGACAGCCCGCTCAAGGCGGG	540
Db	513	CTTCTTCCCGGAGCTGTGACAGTGGCTTGGAGGAGCATGTGACAGCCCGCTCAAGGCGGG	572
Qy	541	AGTGGAGACCAACCAACCTCTCCAAACAGAGCAACAAAGTACGCGGCGCCAGCACTACT	600
Db	573	AGTGGAGACCAACCAACCTCTCCAAACAGAGCAACAAAGTACGCGGCGCCAGCACTACT	632
Qy	601	GAGCCTGAGCGCCCGAGCAAGTGGAAATCCCAAGAAAGTAAAGTGCACAGTCAAGATTA	660
Db	633	GAGCCTGAGCGCCCGAGCAAGTGGAAATCCCAAGAAAGTAAAGTGCACAGTCAAGATTA	692
Qy	661	AGGAGACACCCGTGAGTGAAGACATGTGTC	690
Db	693	AGGAGACACCCGTGAGTGAAGACATGTGTC	721
RESULT 9			
BW920304		1049 bp	mrna
LOCUS			linear
DEFINITION	ABENCOURT 6706549 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5750222		
ACCESSION	BW920304		
VERSION	BW920304.1		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 1049)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabbs-remail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.llnl.gov		
	Plate: L1AM2780		
	row: n		
	column: 21		
	High quality sequence stop: 669.		
FEATURES			
source	Location/Qualifiers		
	1..1049		
	/organism="Homo sapiens"		

```

/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5750228"
/lab_host="DH10B"
/clone_11b="NIH_MGC_120"
/notes="Organ: pooled pancreas and spleen; Vector:
pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."
ORIGIN

Query Match      41.8%; Score 610.8; DB 12; Length 1049;
Best Local Similarity 93.8%; Pred. No. 5,9e-129;
Matches 636; Conservative 0; Mismatches 42; Indels 0; Gaps 0

QY      1  GAGCCGAGGCTCGCATTCCTCAACCTCCCTCACTCAAGGACAGAGGCTCTGGAGCCAGTC 60
Db      166 GAGCTGAGGCTGTCTCTCTCTCAACCCTCTCACTCAAGGACAGAGGCTCTGGAGCCAGTC 225

QY      61  TGCCCTGACTCAGTCGCTCCGATCTCGATCGGATCTCTCGAAGATGATCACCATCTCTG 120
Db      226 TGTCTGACTCAGGCTCCCTCGAGGCTCAGGATCTCTGAGACAGTCACCATCTCTG 285

QY      121 CACTGGAACCGACAGTACGTTGGTGGTTAACTATGTCTCTGGTACCAACAGACCC 180
Db      286 CGCTGGAACCGAGCGTGAAGTTGAAATTAACCAATGCTCTGGTTTCAACAGACCC 345

QY      181 AGGCAAGCCCCCAACTCATCTATTATGAGATCGATATGAGCCCTCAGGGGTTCTTAA 240
Db      346 AGGCAAGCCCCCAACTCATCTATTATGAGATCGATATGAGGCGCCTCAGGGTCCCTGA 405

QY      241 TCGCTTCTTGGCTCCAAAGCTTGGCAACGAGCTCCCTGACCATCTCTGGGCTCAAGC 300
Db      406 TCGCTTCTTGGCTCCAAAGCTTGGCAACGAGCTCCCTGACCATCTCTGGGCTCAAGC 465

QY      301 TGAGGACGAGGCTGATTTTAACTGCTGCTCAATATACAAAGATATTCTCATGCTTCGG 360
Db      466 TGAGGATAGGCTGATTTTAACTGACGCGCATATGACGACAGCAACATATGTCTTCGG 525

QY      466 TGAGGATAGGCTGATTTTAACTGACGCGCATATGACGACAGCAACATATGTCTTCGG 525
Db      526 AACTGAGGACCAAGGTCACCGTCTTAGGTCACCCCAAGGCCCACTGTCACTCTGTT 585

QY      361 AACTGAGGACCAAGGTCACCGTCTTAGGTCACCCCAAGGCCCACTGTCACTCTGTT 420
Db      526 AACTGAGGACCAAGGTCACCGTCTTAGGTCACCCCAAGGCCCACTGTCACTCTGTT 585

QY      421 CCGCGCCCTCTCTAGAGGCTCCAAAGCCAAAGGCCCACTAAGTGTCTGATAGTGA 480
Db      586 CCGCGCCCTCTCTAGAGGCTCCAAAGCCCAAGGCCCACTAAGTGTCTGATAGTGA 645

QY      481 CTTCTACCCCGGAGGTGTGACAGTGCCCTGGAAGGCAAGATGAGCAGCCCGTCAAGCGAG 540
Db      646 CTTCTACCCCGGAGGTGTGACAGTGCCCTGGAAGGCAAGATGAGCAGCCCGTCAAGCGAG 705

QY      541 AGTGAAGACCCACCAACCTCTCCAAACAGAGCAACAACAATGCGGGGCGACGACTACT 600
Db      706 AGTGAAGACCCACCAACCTCTCCAAACAGAGCAACAACAATGCGGGGCGACGACTACT 765

QY      601 GAGCCTGAGCCCGAGCAGTGGAGTCCCAAGAAGCTACAGAGTCCAGTTCACCATGA 660
Db      766 GAGCCTGAGCCCGAGAGCAGTGGAGTCCCAAGAAGCTACAGAGTCCAGTTCACCATGA 825

QY      661 AGGAGACACCGTGAATGA 678
Db      826 AGGAGACACCGTGAATGA 843

RESULT 10
LOCUS       BG759257               686 bp      mRNA      linear      EST 15-MAY-2001
DEFINITION  G02710936F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4851128 5',

```

ACCESSION mRNA sequence.
 BG759257 GI:14069910
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 686)
 NIH-MGC <http://mgi.nci.nih.gov/>,
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LCM1693 row: h column: 09
 High quality sequence stop: 683.
 Location/Qualifiers

FEATURES

Source

1.686
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4851128"
 /tissue_type="Primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_48"
 /note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; CDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACG(C). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match

Best Local Similarity 41.6%; Score 607.8; DB 12; Length 686;
 Pred. No. 2.3e-128;

Matches 624; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

```

1  GGCCTGGAGCTCTGCTATTCCTCAACCTCTCACTGAGGACAGGGTCTGGAGCCAGTC 60
2  |||||
3  34  GGCCTGGAGCTCTGCTATTCCTCAACCTCTCACTGAGGACAGGGTCTGGAGCCAGTC 93
4  |||||
5  61  TGCCTGACTCACTGCTGCTCGGTCTGGGTCTCTGAGCAAGTCATCATCTCTG 120
6  |||||
7  94  TGCCTGACTCACTGCTGCTCGGTCTGGGTCTCTGAGCAAGTCATCATCTCTG 153
8  |||||
9  121  CACTGGAACAGCAGTCAGTGTGTATTAATGTCCTCTGTAACCAACAGACCC 180
10 |||||
11 154  CACTGGAACAGCAGTGTGTATTAATGTCCTCTGTAACCAACAGACCC 213
12 |||||
13 181  AGGCAAGCCCCCAATCATATTATGAGTCAGTAAATCGCCCTCAGGGGTTCTAA 240
14 |||||
15 214  AGGCAAGCCCCCAATCATATTATGAGTCAGTAAATCGCCCTCAGGGGTTCTAA 273
16 |||||
17 241  TGCCTTCTCTGCTTCCAGTCTGGCAACAGGCTCTCCAGCATCTCTGGGCTCCAGGC 300
18 |||||
19 274  TGCCTTCTCTGCTTCCAGTCTGGCAACAGGCTCTCCAGCATCTCTGGGCTCCAGGC 333
20 |||||
21 301  TGAGAGCAGAGCTGATTTATCTGCTCATATACAAAGTACTCTCATGCTCTCG 360
22 |||||
23 334  TGAGAGCAGAGCTGATTTATCTGCTCATATACAAAGTACTCTCATGCTCTCG 393
24 |||||
25 361  AACTGGAGCAAGTCAAGCTCTAGGTCAAGCCCAAGCCCAAGCCCAAGTCTCTGTT 420
26 |||||

```

```

DB 394  AACTGGAGCAAGTCAAGCTCTAGGTCAAGCCCAAGCCCAAGCCCAAGTCTCTGTT 453
QY 421  CCCGCCCTCTCTGAGAGCTCCAGCCCAAGGCCCACTGTGTGTGATCACTGA 480
DB 454  CCCGCCCTCTCTGAGAGCTCCAGCCCAAGGCCCACTGTGTGTGATCACTGA 513
QY 481  CTTTCAACCCGGAGCTGTGACAGTGGCTGGAAGGAGATGAGAGCCCGTCAAGCGGG 540
DB 514  CTTTCAACCCGGAGCTGTGACAGTGGCTGGAAGGAGATGAGAGCCCGTCAAGCGGG 573
QY 541  AGTGAAGACCAACCAACCTCCAAACAGAGCAACACAGTACGCGCCAGCACTACT 600
DB 574  AGTGAAGACCAACCAACCTCCAAACAGAGCAACACAGTACGCGCCAGCACTACT 633
QY 601  GAGCTGACCGCGAGAGTGAAGTCCACAGAGTACAGTCCAGGT 651
DB 634  GAGCTGACCGCGAGAGTGAAGTCCACAGAGTACAGTCCAGGT 684

```

RESULT 11

EX446259

LOCUS

DEFINITION

EX446259 957 bp. mRNA. linear. EST 22-MAY-2003
 Clone CS0D3002YH06 5-PRIME, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 957)
 L.M.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seq@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by life technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 10531.f For
 more information about this cluster, see
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS6AG001ZF020P1&cluster=10531.f>. Contact :
 Feng Liang Email: liang@life.com URL :
<http://fulllength.invitrogen.com/InvitrogenCorporation1600>
 Parady Avenue Genoscope sequence ID : CS6AG001ZF020P1.
 Location/Qualifiers

FEATURES

Source

1.957
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D3002YH06"
 /tissue_type="B CELLS (RAMOS CELL LINE)"
 /cell_line="RAMOS CELL LINE"
 /note="Vector: pOWSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN

Query Match

Best Local Similarity 41.6%; Score 606.8; DB 13; Length 957;
 Pred. No. 4.6e-128;

Matches 653; Conservative 0; Mismatches 37; Indels 4; Gaps 2;

```

1  GGCCTGGAGCTCTGCTATTCCTCAACCTCTCACTGAGGACAGGGTCTGGAGCCAGTC 60
2  |||||
3  97  GGCCTGGAGCTCTGCTATTCCTCAACCTCTCACTGAGGACAGGGTCTGGAGCCAGTC 156
4  |||||
5  61  TGCCTGACTCACTGCTGCTCGGTCTGGGTCTCTGAGCAAGTCATCATCTCTGTT 120
6  |||||

```

[illegible]

VERSION BX462166.1 GI:31033379
 EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1201)
 AUTHORS Li, W.B., Gruber, C., Jessee, V., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 10531.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DG002D03QPI&cluster=10531.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS0DG002D03QPI.
 Location/Qualifiers
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DG002YH06"
 /rname="B CELLS (RAMOS CELL LINE)"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
 /note="Vector: PCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo (dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the PCMVSPORT 6 vector.
 Library was not normalized."

FEATURES
 source

ORIGIN

Query Match 41.5%; Score 606.4; DB 13; Length 1201;
 Best Local Similarity 93.9%; Pred. No. 6,4e-128;
 Matches 652; Conservative 1; Mismatches 37; Indels 4; Gaps 2;

```

QY 1 GGCCTGGGCTCTGCTATTCCTCACTCCTCACTCAGGCAAGGCTCTGGGCCAAGC 60
DB 86 GGCCTGGGCTCTGCTATTCCTCACTCCTCACTCAGGCAAGGCTCTGGGCCAAGC 145
QY 61 TGCCTGACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB 146 TGCCTGACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 205
QY 121 CACTGGAACGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB 206 CACTGGAACGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 265
QY 181 AGGCAAGCCCACTCACTATTCATGAGTCACTAAATGCGCCCTCAGGAGTTCTAA 240
DB 266 AGGCAAGCCCACTCACTATTCATGAGTCACTAAATGCGCCCTCAGGAGTTCTAA 325
QY 241 TCGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 326 TCGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 385
QY 301 TGAAGACGAGGCTGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 357
DB 386 TGAAGACGAGGCTGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 445
QY 358 CGGAATCTGGAACCAAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 417
DB 446 CGGAATCTGGAACCAAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 505
QY 418 GTTCCGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 477
DB 506 GTTCCGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 565

```

```

QY 478 TGAATTCACCCGGAGCTGACAGTGGCTGGAGGACGATGGACGCCCCCTCAAGGC 537
DB 566 TGAATTCACCCGGAGCTGACAGTGGCTGGAGGACGATGGACGCCCCCTCAAGGC 625
QY 538 GGAAGTGAAGCAACCAACCTCTCAACAGAGCAACAGTACGCGGCCAGACCTA 597
DB 626 GGAAGTGAAGCAACCAACCTCTCAACAGAGCAACAGTACGCGGCCAGACCTA 685
QY 598 CCTGAGCTGACGCTGAGGAGTGAAGTCCACAGAGCTACAGTCCAGAGTCAAGCA 657
DB 686 CCTGAGCTGACGCTGAGGAGTGAAGTCCACAGAGCTACAGTCCAGAGTCAAGCA 745
QY 658 TGAAGGAGCAACCTGATGATGAGACAGTGTCC 691
DB 746 TGAAGGAGCAACCTGATGATGAGACAGTGTCC 778

```

RESULT 14

EX442894 1201 bp mRNA linear EST 15-MAY-2003
 LOCUS BX442894 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
 DEFINITION clone CS0DG001YK17 5-PRIME, mRNA sequence.
 ACCESSION BX442894
 VERSION BX442894.1 GI:30772087
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1201)
 AUTHORS Li, W.B., Gruber, C., Jessee, V., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 10531.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DG001AF09QPI&cluster=10531.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS0DG001AF09QPI.
 Location/Qualifiers
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DG001YK17"
 /rname="B CELLS (RAMOS CELL LINE)"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
 /note="Vector: PCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo (dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the PCMVSPORT 6 vector.
 Library was not normalized."

FEATURES

source

ORIGIN

Query Match 41.4%; Score 604; DB 13; Length 1201;
 Best Local Similarity 93.5%; Pred. No. 2,3e-127;
 Matches 649; Conservative 2; Mismatches 39; Indels 4; Gaps 2;

```

QY 1 GGCCTGGGCTCTGCTATTCCTCACTCCTCACTCAGGCAAGGCTCTGGGCCAAGC 60
DB 100 GGCCTGGGCTCTGCTATTCCTCACTCCTCACTCAGGCAAGGCTCTGGGCCAAGC 159
QY 61 TGCCTGACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB 160 TGCCTGACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 219

```


QY 121 CACTGAAACGACGATCGTGGTATATATATGTCCTGGTACCAAGCAGACCC 180
 Db 220 CACTGAAACGACGATCGTGGTATATATATGTCCTGGTACCAAGCAGACCC 279
 QY 181 AGCGAAAGCCCCCAACTCATATTTATAGAGTCAATTCGCGCCCTCAGGGGTTCTAA 240
 Db 280 AGCGAAAGCCCCCAACTCATATTTATAGAGTCAATTCGCGCCCTCAGGGGTTCTAA 339
 QY 241 TCGCTCTCTGCTCCCAAGTGTGCAACAGCGCTCCCTGACCTCTCTGCTCCAGGC 300
 Db 340 TCGCTCTCTGCTCCCAAGTGTGCAACAGCGCTCCCTGACCTCTCTGCTCCAGGC 399
 QY 301 TGAGGACGAGGCTATTTATCTGCTCTCATATACAA---GAAGTACTTCTCATGCTT 357
 Db 400 TGAGGACGAGGCTATTTATCTGCTCTCATATACAAAGAGACGCAATTCAGAGATT 459
 QY 358 CGGAACTGGAGCAAGGTCAACCTGCTTGAAGTCAAGCCCAAGCCCACTGCTACTCT 417
 Db 460 CGGAGAGGAGCAAGTCAACCTGCTTGAAGTCAAGCCCAAGCCCACTGCTACTCT 519
 QY 418 GTTCCCGCCCTCTCTGAGAGTCAAGCCCAAGCCCACTGCTGCTGATCAG 477
 Db 520 GTTCCCGCCCTCTCTGAGAGTCAAGCCCAAGCCCACTGCTGCTGATCAG 579
 QY 478 TGACTTCAACCGGAGCTGTGACAGTGGCTGGAGGCAATGGACGCCGCTCAAGGC 537
 Db 580 TKACTTCAACCGGAGCTGTGACAGTGGCTGGAGGCAATGGACGCCGCTCAAGGC 639
 QY 538 GGGAGTGGAGCAACCAACCTCCAAACAGAGCAAGCAAGTACCGCGCCGACAGCTA 557
 Db 640 GGGAGTGGAGCAACCAACCTCCAAACAGAGCAAGTACCGCGCCGACAGCTA 659
 QY 598 CCTGAGCTGAGCGCCGACAGTGAAGTCCACAGAGTCAAGCTGACAGCTCAAGCA 657
 Db 700 CCTGAGCTGAGCGCCGACAGTGAAGTCCACAGAGTCAAGCTGACAGCTCAAGCA 759
 QY 658 TGAAGGAGCACCGTGAATGAAGACAGTGTCCC 691
 Db 760 TGAAGGAGCACCGTGAATGAAGACAGTGTCCC 792

RESULT 15
 LOCUS BG397282 840 bp mRNA linear EST 12-MAR-2001
 DEFINITION 60243065F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4565272 5',
 mRNA sequence.
 ACCESSION BG397282
 VERSION BG397282.1 GI:13290730
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 840)
 NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rgs@dbp-remail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: L1CM1282 row: a column: 17
 High quality sequence stop: 840.
 Location/Qualifiers
 1..840
 /organism="Homo sapiens"
 /mol_type="mRNA"

FEATURES
 SOURCE

/db xref="taxon:9606"
 /clone="IMAGE:4565272"
 /tissue_type="primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_11b="NIH_MGC_48"
 /note="Organ: B-cells; Vector: pOT87; Site: 1: XhoI;
 Site: 2: EcoRI; cDNA made by oligo-dT priming.
 directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG (G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong, in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using 2AP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

Query Match 41.4%; Score 603.8; DB 12; Length 840;
 Best Local Similarity 95.6%; Pred. No. 2.1e-127;
 Matches 653; Conservative 0; Mismatches 27; Indels 3; Gaps 3;
 QY 8 GCTGTGCTATTCTCTACCCCTCTCTCACTCAGGAGCAGAGGCTCTGAGGCTGCTG 57
 Db 1 GCTGTGCTGCTCTCTACCCCTCTCTCACTCAGGAGCAGAGGCTCTGAGGCTGCTG 60
 QY 68 ACTCAGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 127
 Db 61 ACTCAGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 QY 128 ACCAGAGTCAAGTGGTGGTATTAATGATGCTGCTGCTGCTGCTGCTGCTGCTG 187
 Db 121 ACCAGAGTCAAGTGGTGGTATTAATGATGCTGCTGCTGCTGCTGCTGCTGCTG 180
 QY 188 GCCCCCAACTCATATTTATGAGTCAATGAGGCTGAGGCTGAGGCTGCTGCTGCTG 247
 Db 181 GCCCCCAACTCATATTTATGAGTCAATGAGGCTGAGGCTGAGGCTGCTGCTGCTG 240
 QY 248 TTGAGCTCAAGTCTGAGCAACGAGCTCTCTGACATCTCTGAGCTCTGAGCTGAG 307
 Db 241 TTGAGCTCAAGTCTGAGCAACGAGCTCTCTGACATCTCTGAGCTCTGAGCTGAG 300
 QY 308 GAGGCTGATTTATTTAGTCTGCTCATATACAAAGATCTTCTCATGCTCTTGGAACTG 367
 Db 301 GAGGCTGATTTATTTAGTCTGCTCATATACAAAGATCTTCTCATGCTCTTGGAACTG 360
 QY 368 ACCAAGTCAACCGTCTGAGTCAAGCCCAAGGCCAACCCCACTGCTCTGCTGCTGCTG 427
 Db 361 ACCAAGTCAACCGTCTGAGTCAAGCCCAAGGCCAACCCCACTGCTCTGCTGCTGCTG 420
 QY 428 TTCTCTGAGAGCTTCAAGCCCAAGGCCCAACAGTGTGTGATCAAGTGTCTTAC 487
 Db 421 TTCTCTGAGAGCTTCAAGCCCAAGGCCCAACAGTGTGTGATCAAGTGTCTTAC 480
 QY 488 CCGGAGTGTGACAGTGGCTGTGAGGAGCAGATGGACCCCGTCAAGGCGGAGTGGAG 547
 Db 481 CCGGAGTGTGACAGTGGCTGTGAGGAGCAGATGGACCCCGTCAAGGCGGAGTGGAG 540
 QY 548 ACCACCAAACTCTCCAAACAGAGCAACAAAGTACGAGGCTGAGTCTGAGCTGCTG 607
 Db 541 ACCACCAAACTCTCCAAACAGAGCAACAAAGTACGAGGCTGAGTCTGAGCTGCTG 598
 QY 608 ACCGCGAGCAGTGAATGCCCAAGAGTCAAGTCTGAGTCAAGTCAAGTCAAGTCAAG 667
 Db 599 ACCGCGAGCAGTGAATGCCCAAGAGTCAAGTCTGAGTCAAGTCAAGTCAAGTCAAG 658
 QY 668 ACCGTGATGAAGACAGTGTCC 690
 Db 659 ACCGTGATGAAGACAGTGTCC 680

Search completed: April 26, 2004, 20:25:06
 Job time : 3554.63 secs